

POLYNUCLEOTIDES ENCODING A NOVEL  
HUMAN PHOSPHATASE, BMY\_HPP13

This application claims benefit to provisional application U.S. Serial No.  
5 60/393,253 filed July 2, 2002, under 35 U.S.C. 119(e). The entire teachings of the  
referenced application are incorporated herein by reference.

FIELD OF THE INVENTION

The present invention provides novel polynucleotides encoding a human  
10 phosphatase polypeptide, BMY\_HPP13, fragments and homologues thereof. Also  
provided are vectors, host cells, antibodies, and recombinant and synthetic methods  
for producing said polypeptide. The invention further relates to diagnostic and  
therapeutic methods for applying this novel human phosphatase polypeptide to the  
diagnosis, treatment, and/or prevention of various diseases and/or disorders related to  
15 these polypeptides. The invention further relates to screening methods for identifying  
agonists and antagonists of the polynucleotides and polypeptides of the present  
invention.

BACKGROUND OF THE INVENTION

20 Phosphorylation of proteins is a fundamental mechanism for regulating  
diverse cellular processes. While the majority of protein phosphorylation occurs at  
serine and threonine residues, phosphorylation at tyrosine residues is attracting a great  
deal of interest since the discovery that many oncogene products and growth factor  
receptors possess intrinsic protein tyrosine kinase activity. The importance of protein  
25 tyrosine phosphorylation in growth factor signal transduction, cell cycle progression  
and neoplastic transformation is now well established (Hunter et al., Ann. Rev.  
Biochem. 54:987-930 (1985), Ullrich et al., Cell 61:203-212 (1990), Nurse, Nature  
344:503-508 (1990), Cantley et al, Cell 64:281-302 (1991)).

Biochemical studies have shown that phosphorylation on tyrosine residues of a  
30 variety of cellular proteins is a dynamic process involving competing phosphorylation  
and dephosphorylation reactions. The regulation of protein tyrosine phosphorylation  
is mediated by the reciprocal actions of protein tyrosine kinases (PTKases) and



protein tyrosine phosphatases (PTPases). The tyrosine phosphorylation reactions are catalyzed by PTKases. Tyrosine phosphorylated proteins can be specifically dephosphorylated through the action of PTPases. The level of protein tyrosine phosphorylation of intracellular substances is determined by the balance of PTKase and PTPase activities. (Hunter, T., *Cell* 58:1013-1016 (1989)).

The protein tyrosine kinases (PTKases) are a large family of proteins that includes many growth factor receptors and potential oncogenes. (Hanks et al., *Science* 241:42-52 (1988)). Many PTKases have been linked to initial signals required for induction of the cell cycle (Weaver et al., *Mol. Cell. Biol.* 11, 9:4415-4422 (1991)). PTKases comprise a discrete family of enzymes having common ancestry with, but major differences from, serine/threonine-specific protein kinases (Hanks et al., *supra*). The mechanisms leading to changes in activity of PTKases are best understood in the case of receptor-type PTKases having a transmembrane topology (Ullrich et al. (1990) *supra*). The binding of specific ligands to the extracellular domain of members of receptor-type PTKases is thought to induce their oligomerization leading to an increase in tyrosine kinase activity and activation of the signal transduction pathways (Ullrich et al., (1990) *supra*). Deregulation of kinase activity through mutation or overexpression is a well established mechanism for cell transformation (Hunter et al., (1985) *supra*; Ullrich et al., (1990) *supra*).

The protein phosphatases are composed of at least two separate and distinct families (Hunter, T. (1989) *supra*) the protein serine/threonine phosphatases and the protein tyrosine phosphatases (PTPases).

The protein tyrosine phosphatases (PTPases) are a family of proteins that have been classified into two subgroups. The first subgroup is made up of the low molecular weight, intracellular enzymes that contain a single conserved catalytic phosphatase domain. All known intracellular type PTPases contain a single conserved catalytic phosphatase domain. Examples of the first group of PTPases include (1) placental PTPase 1B (Charbonneau et al., *Proc. Natl. Acad. Sci. USA* 86:5252-5256 (1989); Chernoff et al., *Proc. Natl. Acad. Sci. USA* 87:2735-2789 (1989)), (2) T-cell PTPase (Cool et al., *Proc. Natl. Acad. Sci. USA* 86:5257-5261 (1989)), (3) rat brain PTPase (Guan et al., *Proc. Natl. Acad. Sci. USA* 87:1501-1502 (1990)), (4) neuronal phosphatase (STEP) (Lombroso et al., *Proc. Natl. Acad. Sci. USA* 88:7242-7246

(1991)), and (5) cytoplasmic phosphatases that contain a region of homology to cytoskeletal proteins (Gu et al., *Proc. Natl. Acad. Sci. USA* 88:5867-57871 (1991); Yang et al., *Proc. Natl. Acad. Sci. USA* 88:5949-5953 (1991)).

Enzymes of this class are characterized by an active site motif of CX<sub>5</sub>R.  
 5 Within this motif the Cysteine sulfur acts as a nucleophile which cleaves the P-O bond and releases the phosphate; the Arginine interacts with the phosphate and facilitates nucleophilic attack. In many cases the Cysteine is preceded by a Histidine and the Arginine is followed by a Serine or Threonine. In addition, an Aspartate residue located 20 or more amino acids N terminal to the Cysteine acts as a general acid  
 10 during cleavage [Fauman, 1996].

The second subgroup of protein tyrosine phosphatases is made up of the high molecular weight, receptor-linked PTPases, termed R-PTPases. R-PTPases consist of a) an intracellular catalytic region, b) a single transmembrane segment, and c) a putative ligand-binding extracellular domain (Gebbink et al., *supra*).

15 The structures and sizes of the putative ligand-binding extracellular "receptor" domains of R-PTPases are quite divergent. In contrast, the intracellular catalytic regions of R-PTPases are highly homologous. All RPTPases have two tandemly duplicated catalytic phosphatase homology domains, with the prominent exception of an R-PTPase termed HPTP.beta., which has "only one catalytic phosphatase domain.  
 20 (Tsai et al., *J. Biol. Chem.* 266(16):10534-10543 (1991)).

One example of R-PTPases are the leukocyte common antigens (LCA) (Ralph, S. J., *EMBO J.* 6:1251-1257 (1987)). LCA is a family of high molecular weight glycoproteins expressed on the surface of all leukocytes and their hemopoietic progenitors (Thomas, *Ann. Rev. Immunol.* 7:339-369 (1989)). A remarkable degree  
 25 of similarity is detected with the sequence of LCA from several species (Charbonneau et al., *Proc. Natl. Acad. Sci. USA* 85:7182-7186 (1988)). LCA is referred to in the literature by different names, including T200 (Trowbridge et al., *Eur. J. Immunol.* 6:557-562 (1962)), B220 for the B cell form (Coffman et al., *Nature* 289:681-683 (1981)), the mouse allotypic marker Ly-5 (Komuro et al., *Immunogenetics* 1:452-456 (1975)), and more recently CD45 (Cobbald et al., *Leucocyte Typing III*, ed. A. J. McMichael et al., pp. 788-803 (1987)).  
 30

Several studies suggest that CD45 plays a critical role in T cell activation. These studies are reviewed in Weiss A., *Ann. Rev. Genet.* 25:487-510 (1991). In one study, T-cell clones that were mutagenized by NSG and selected for their failure to express CD45 had impaired responses to T-cell receptor stimuli (Weaver et al., (1991) supra). These T-cell clones were functionally defective in their responses to signals transmitted through the T cell antigen receptor, including cytolysis of appropriate targets, proliferation, and lymphokine production (Weaver et al., (1991) supra).

Other studies indicate that the PTPase activity of CD45 plays a role in the activation of pp56.sup.lck, a lymphocyte-specific PTKase (Mustelin et al., *Proc. Natl. Acad. Sci. USA* 86:6302-6306 (1989); Ostergaard et al., *Proc. Natl. Acad. Sci. USA* 86:8959-8963 (1989)). These authors hypothesized that the phosphatase activity of CD45 activates pp56.sup.lck by dephosphorylation of a C-terminal tyrosine residue, which may, in turn, be related to T-cell activation.

Another example of R-PTPases is the leukocyte common antigen related molecule (LAR) (Streuli et al., *J. Exp. Med.* 168:1523-1530 (1988)). LAR was initially identified as a homologue of LCA (Streuli et al., supra). Although the a) intracellular catalytic region of the LAR molecule contains two catalytic phosphatase homology domains (domain I and domain II), mutational analyses suggest that only domain I has catalytic phosphatase activity, whereas domain II is enzymatically inactive (Streuli et al., *EMBO J.* 9(8):2399-2407 (1990)). Chemically induced LAR mutants having tyrosine at amino acid position 1379 changed to a phenylalanine are temperature-sensitive (Tsai et al., *J. Biol. Chem.* 266(16):10534-10543 (1991)).

A new mouse R-PTP, designated mRPTP.mu., has been cloned which has an extracellular domain that shares some structural motifs with LAR. (Gebbink et al., (1991) supra). In addition, these authors have cloned the human homologue of RPTP.mu. and localized the gene on human chromosome 18.

Two *Drosophila* PTPases, termed DLAR and DPTP, have been predicted based on the sequences of cDNA clones (Streuli et al., *Proc. Natl. Acad. Sci. USA* 86:8698-8702 (1989)). cDNAs coding for another *Drosophila* R-PTPase, termed DPTP 99A, have been cloned and characterized (Hariharan et al., *Proc. Natl. Acad. Sci. USA* 88:11266-11270 (1991)).

Other examples of R-PTPases include R-PTPase-.alpha., .beta., .gamma., and .zeta. (Krueger et al., EMBO J. 9:3241-3252 (1990), Sap et al., Proc. Natl. Acad. Sci. USA 87:6112-6116 (1990), Kaplan et al., Proc. Natl. Acad. Sci. USA 87:7000-7004 (1990), Jirik et al., FEBS Lett. 273:239-242 (1990); Mathews et al., Proc. Natl. Acad. Sci. USA 87:4444-4448 (1990), Ohagi et al., Nucl. Acids Res. 18:7159 (1990)).

5 Published application W092/01050 discloses human R-PTPase-.alpha., .beta. and .gamma., and reports on the nature of the structural homologies found among the conserved domains of these three R-PTPases and other members of this protein family. The murine R-PTPase-.alpha. has 794 amino acids, whereas the human R-

10 PTPase-.alpha. has 802 amino acids. R-PTPase-.alpha. has an intracellular domain homologous to the catalytic domains of other tyrosine phosphatases. The 142 amino acid extracellular domain (including signal peptide of RPTPase-.alpha.) has a high serine and threonine content (32%) and 8 potential N-glycosylation sites. cDNA clones have been produced that code for the R-PTPase-.alpha., and R-PTPase-.alpha.

15 has been expressed from eukaryotic hosts. Northern analysis has been used to identify the natural expression of R-PTPase-.alpha. in various cells and tissues. A polyclonal antibody to R-PTPase-.alpha. has been produced by immunization with a synthetic peptide of R-PTPase-.alpha., which identifies a 130 kDa protein in cells transfected with a cDNA clone encoding a portion of R-PTPase-.alpha..

20 Another example of R-PTPases is HePTP. (Jirik et al, FASEB J. 4:82082 (1990) Abstract 2253). Jirik et al. screened a cDNA library derived from a hepatoblastoma cell line, HepG2, with a probe encoding the two PTPase domains of LCA, and discovered a cDNA clone encoding a new RPTPase, named HePTP. The HePTP gene appeared to be expressed in a variety of human and murine cell lines and

25 tissues.

Since the initial purification, sequencing, and cloning of a PTPase, additional potential PTPases have been identified at a rapid pace. The number of different PTPases that have been identified is increasing steadily, leading to speculations that this family may be as large as the PTKase family (Hunter (1989) supra).

30 Conserved amino acid sequences in the catalytic domains of known PTPases have been identified and defined (Krueger et al., EMBO J. 9:3241-3252 (1990) and Yi

et al., Mol. Cell. Biol. 12:836-846 (1992), which are incorporated herein by reference.) These amino acid sequences are designated "consensus sequences" herein.

Yi et al. aligned the catalytic phosphatase domain sequences of the following PTPases: LCA, PTP1B, TCPTP, LAR, DLAR, and HPTP.alpha., HPTP.beta., and  
 5 HPTP.gamma.. This alignment includes the following "consensus sequences" (Yi et al., supra, FIG. 2(A), lines 1 and 2): DYINAS/N (SEQ ID NO:5), CXXYWP (SEQ ID NO:6), and I/VVMXXXXXE (SEQ ID NO:7).

Krueger et al., aligned the catalytic phosphatase domain sequences of PTP1B, TCPTP, LAR, LCA, HPTP.alpha., .beta., .gamma., .GAMMA., .delta., .epsilon. and  
 10 .zeta. and DLAR and DPTP. This alignment includes the following "consensus sequences: (Krueger et al., supra, FIG. 7, lines 1 and 2): D/NYINAS/N (SEQ ID NO:8), CXXYWP (SEQ ID NO:9), and I/VVMXXXXXE (SEQ ID NO:10).

It is becoming clear that dephosphorylation of tyrosine residues can by itself function as an important regulatory mechanism. Dephosphorylation of a C-terminal  
 15 tyrosine residue has been shown to activate tyrosine kinase activity in the case of the src family of tyrosine kinases (Hunter, T. Cell 49:1-4 (1987)). Tyrosine dephosphorylation has been suggested to be an obligatory step in the mitotic activation of the maturation-promoting factor (MPF) kinase (Morla et al., Cell 58:193-203 (1989)). These observations point out the need in the art for  
 20 understanding the mechanisms that regulate tyrosine phosphatase activity.

Modulators (inhibitors or activators) of human phosphatase expression or activity could be used to treat a subject with a disorder characterized by aberrant phosphatase expression or activity or by decreased phosphorylation of a phosphatase substrate protein. Examples of such disorders include but are not limited to: an  
 25 immune, anti-proliferative, proliferative (e.g. cancer), metabolic (e.g. diabetes or obesity), bone (e.g., osteoporosis), neural, and/or cardiovascular diseases and/or disorders, in addition to, viral pathogenesis.

It is clear that further analysis of structure-function relationships among PTPases are needed to gain important understanding of the mechanisms of signal  
 30 transduction, cell cycle progression and cell growth, and neoplastic transformation.

The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing

the recombinant vectors, as well as to methods of making such vectors and host cells, in addition to their use in the production of human phosphatase polypeptides or peptides using recombinant techniques. Synthetic methods for producing the polypeptides and polynucleotides of the present invention are provided. Also provided are diagnostic methods for detecting diseases, disorders, and/or conditions related to the human phosphatase polypeptides and polynucleotides, and therapeutic methods for treating such diseases, disorders, and/or conditions. The invention further relates to screening methods for identifying binding partners of the polypeptides.

## BRIEF SUMMARY OF THE INVENTION

The present invention provides isolated nucleic acid molecules, that comprise, or alternatively consist of, a polynucleotide encoding the human BMY\_HPP13 phosphatase protein having the amino acid sequence shown as SEQ ID NO:2, or the amino acid sequence encoded by the cDNA clone, BMY\_HPP13, deposited as ATCC Deposit Number PTA-4803 on November 14, 2002.

The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells, in addition to their use in the production of human phosphatase polypeptides or peptides using recombinant techniques. Synthetic methods for producing the polypeptides and polynucleotides of the present invention are provided. Also provided are diagnostic methods for detecting diseases, disorders, and/or conditions related to the human phosphatase polypeptides and polynucleotides, and therapeutic methods for treating such diseases, disorders, and/or conditions. The invention further relates to screening methods for identifying binding partners of the polypeptides.

The invention further provides an isolated BMY\_HPP13 human phosphatase polypeptide having an amino acid sequence encoded by a polynucleotide described herein.

The invention further relates to a polynucleotide encoding a polypeptide fragment of SEQ ID NO:2, or a polypeptide fragment encoded by the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1.

The invention further relates to a polynucleotide encoding a polypeptide domain of SEQ ID NO:2 or a polypeptide domain encoded by the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1.

5 The invention further relates to a polynucleotide encoding a polypeptide epitope of SEQ ID NO:2 or a polypeptide epitope encoded by the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1.

The invention further relates to a polynucleotide encoding a polypeptide of SEQ ID NO:2 or the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1, having biological activity.

10 The invention further relates to a polynucleotide which is a variant of SEQ ID NO:1.

The invention further relates to a polynucleotide which is an allelic variant of SEQ ID NO:1.

15 The invention further relates to a polynucleotide which encodes a species homologue of the SEQ ID NO:2.

The invention further relates to a polynucleotide which represents the complimentary sequence (antisense) of SEQ ID NO:1.

20 The invention further relates to a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified herein, wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

The invention further relates to an isolated nucleic acid molecule of SEQ ID NO:2, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a human phosphatase protein.

25 The invention further relates to an isolated nucleic acid molecule of SEQ ID NO:1 wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:2 or the polypeptide encoded by the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1.

30 The invention further relates to an isolated nucleic acid molecule of of SEQ ID NO:1, wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:1 or the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1.

The invention further relates to an isolated nucleic acid molecule of SEQ ID NO:1, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

5 The invention further relates to an isolated polypeptide comprising an amino acid sequence that comprises a polypeptide fragment of SEQ ID NO:2 or the encoded sequence included in the deposited clone.

The invention further relates to a polypeptide fragment of SEQ ID NO:2 or the encoded sequence included in the deposited clone, having biological activity.

10 The invention further relates to a polypeptide domain of SEQ ID NO:2 or the encoded sequence included in the deposited clone.

The invention further relates to a polypeptide epitope of SEQ ID NO:2 or the encoded sequence included in the deposited clone.

The invention further relates to a full length protein of SEQ ID NO:2 or the encoded sequence included in the deposited clone.

15 The invention further relates to a variant of SEQ ID NO:2.

The invention further relates to an allelic variant of SEQ ID NO:2.

The invention further relates to a species homologue of SEQ ID NO:2.

20 The invention further relates to the isolated polypeptide of SEQ ID NO:2, wherein the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

The invention further relates to an isolated antibody that binds specifically to the isolated polypeptide of SEQ ID NO:2.

25 The invention further relates to a method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of SEQ ID NO:2 or the polynucleotide of SEQ ID NO:1.

30 The invention further relates to a method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising the steps of (a) determining the presence or absence of a mutation in the polynucleotide of SEQ ID NO:1; and (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

The invention further relates to a method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising the steps of (a) determining the presence or amount of expression of the polypeptide of SEQ ID NO:2 in a biological sample; and (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

The invention further relates to a method for identifying a binding partner to the polypeptide of SEQ ID NO:2 comprising the steps of (a) contacting the polypeptide of SEQ ID NO:2 with a binding partner; and (b) determining whether the binding partner effects an activity of the polypeptide.

The invention further relates to a gene corresponding to the cDNA sequence of SEQ ID NO:1.

The invention further relates to a method of identifying an activity in a biological assay, wherein the method comprises the steps of (a) expressing SEQ ID NO:1 in a cell, (b) isolating the supernatant; (c) detecting an activity in a biological assay; and (d) identifying the protein in the supernatant having the activity.

The invention further relates to a process for making polynucleotide sequences encoding gene products having altered activity selected from the group consisting of SEQ ID NO:2 activity comprising the steps of (a) shuffling a nucleotide sequence of SEQ ID NO:1, (b) expressing the resulting shuffled nucleotide sequences and, (c) selecting for altered activity selected from the group consisting of SEQ ID NO:2 activity as compared to the activity selected from the group consisting of SEQ ID NO:2 activity of the gene product of said unmodified nucleotide sequence.

The invention further relates to a shuffled polynucleotide sequence produced by a shuffling process, wherein said shuffled DNA molecule encodes a gene product having enhanced tolerance to an inhibitor of any one of the activities selected from the group consisting of SEQ ID NO:2 activity.

The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a condition related to aberrant phosphatase activity.

The invention further relates to a method of identifying a compound that modulates the biological activity of a phosphatase, comprising the steps of, (a) combining a candidate modulator compound with a phosphatase having the sequence set forth in one or more of SEQ ID NO:2; and (b) measuring an effect of the candidate modulator compound on the activity of a phosphatase.

The invention further relates to a method of identifying a compound that modulates the biological activity of a phosphatase, comprising the steps of, (a) combining a candidate modulator compound with a host cell expressing a phosphatase having the sequence as set forth in SEQ ID NO:2; and , (b) measuring an effect of the candidate modulator compound on the activity of the expressed a phosphatase.

The invention further relates to a method of identifying a compound that modulates the biological activity of a phosphatase, comprising the steps of, (a) combining a candidate modulator compound with a host cell containing a vector described herein, wherein a phosphatase is expressed by the cell; and, (b) measuring an effect of the candidate modulator compound on the activity of the expressed a phosphatase.

The invention further relates to a method of screening for a compound that is capable of modulating the biological activity of a phosphatase, comprising the steps of: (a) providing a host cell described herein; (b) determining the biological activity of a phosphatase in the absence of a modulator compound; (c) contacting the cell with the modulator compound; and (d) determining the biological activity of a phosphatase in the presence of the modulator compound; wherein a difference between the activity of a phosphatase in the presence of the modulator compound and in the absence of the modulator compound indicates a modulating effect of the compound.

The invention further relates to a compound that modulates the biological activity of human a phosphatase as identified by the methods described herein.

The invention further relates to a method for preventing, treating, or ameliorating a medical condition, wherein the medical condition is a immune condition.

The invention further relates to a method for preventing, treating, or ameliorating a medical condition, wherein the medical condition is an inflammatory disease.

The invention further relates to a method for preventing, treating, or ameliorating a medical condition, wherein the medical condition is an inflammatory disease where dual-specificity phosphatases, either directly or indirectly, are involved in disease progression.

5       The invention further relates to a method for preventing, treating, or ameliorating a medical condition, wherein the medical condition is a cancer.

The invention further relates to a method for preventing, treating, or ameliorating a medical condition, wherein the medical condition is a neural disorder.

10       The invention further relates to a method for preventing, treating, or ameliorating a medical condition, wherein the medical condition is a reproductive disorder.

The invention further relates to a method for preventing, treating, or ameliorating a medical condition, wherein the medical condition is an gastrointestinal disorder.

15       The invention further relates to a method for preventing, treating, or ameliorating a medical condition, wherein the medical condition is a hepatic disorder.

The invention further relates to a method for preventing, treating, or ameliorating a medical condition, wherein the medical condition is an endocrine disorder.

20       The invention further relates to a method for preventing, treating, or ameliorating a medical condition, wherein the medical condition is a pulmonary disorder.

25       The present invention also provides structure coordinates of the homology model of the BMY\_HPP13 polypeptide (SEQ ID NO:2) provided in Figure 8. The complete coordinates are listed in Table IV. The model of the present invention further provide a basis for designing stimulators and inhibitors or antagonists of one or more of the biological functions of BMY\_HPP13, or of mutants with altered ligand binding specificity.

30       The invention also provides a machine readable storage medium which comprises the structure coordinates of BMY\_HPP13, including all or any parts conserved calpain regions. Such storage medium encoded with these data are capable of displaying on a computer screen or similar viewing device, a three-dimensional

graphical representation of a molecule or molecular complex which comprises said regions or similarly shaped homologous regions.

5 The invention also provides a machine-readable data storage medium, comprising a data storage material encoded with machine readable data, wherein the data is defined by the structure coordinates of the model BMY\_HPP13 according to Table IV or a homologue of said model, wherein said homologue comprises any kind of surrogate atoms that have a root mean square deviation from the backbone atoms of the complex of not more than about 4.0, 3.0, 2.0, 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, or 0.1 Angstroms.

10 The invention also provides a model comprising all or any part of the model defined by structure coordinates of BMY\_HPP13 according to Table IV, or a mutant or homologue of said molecule or molecular complex.

The invention also provides a method for identifying a mutant of BMY\_HPP13 with altered biological properties, function, or reactivity, the method  
15 comprising one or more of the following steps: (a) use of the model or a homologue of said model according to Table IV, for the design of protein mutants with altered biological function or properties which exhibit any combination of therapeutic effects described herein; and/or (b) use of the model or a homologue of said model, for the design of a protein with mutations in the ligand binding site region comprised of the  
20 amino acids Y45, Q47, R48, R87, D150 N197, Q198, A199, K200, N201, Q202, and/or S203 of SEQ ID NO:2 according to Table IV with altered biological function or properties which exhibit any combination of therapeutic effects described herein.

The method also relates to a method for identifying modulators of BMY\_HPP13 biological properties, function, or reactivity, the method comprising the  
25 step of modeling test compounds that fit spatially into the active site region defined by all or any portion of residues Y45, Q47, R48, R87, D150 N197, Q198, A199, K200, N201, Q202, and/or S203 of the three-dimensional structural model according to Table IV, or using a homologue or portion thereof, or analogue in which the original C, N, and O atoms have been replaced with other elements

30 The invention also provides methods for designing, evaluating and identifying compounds which bind to all or parts of the aforementioned regions. The methods include three dimensional model building (homology modeling) and methods of

computer assisted-drug design which can be used to identify compounds which bind or modulate the forementioned regions of the BMY\_HPP13 polypeptide. Such compounds are potential inhibitors of BMY\_HPP13 or its homologues.

The invention also relates to a method of using said structure coordinates as set forth in Table IV to identify structural and chemical features of BMY\_HPP13; employing identified structural or chemical features to design or select compounds as potential BMY\_HPP13 modulators; employing the three-dimensional structural model to design or select compounds as potential BMY\_HPP13 modulators; synthesizing the potential BMY\_HPP13 modulators; screening the potential BMY\_HPP13 modulators in an assay characterized by binding of a protein to the BMY\_HPP13. The invention also relates to said method wherein the potential BMY\_HPP13 modulator is selected from a database. The invention further relates to said method wherein the potential BMY\_HPP13 modulator is designed de novo. The invention further relates to a method wherein the potential BMY\_HPP13 modulator is designed from a known modulator of activity.

#### BRIEF DESCRIPTION OF THE FIGURES/DRAWINGS

Figures 1A-B show the polynucleotide sequence (SEQ ID NO:1) and deduced amino acid sequence (SEQ ID NO:2) of the novel full-length human dual specificity phosphatase, BMY\_HPP13, of the present invention. The standard one-letter abbreviation for amino acids is used to illustrate the deduced amino acid sequence. The polynucleotide sequence of BMY\_HPP13 contains a sequence of 989 nucleotides (SEQ ID NO:1), encoding a polypeptide of 246 amino acids (SEQ ID NO:2). An analysis of the BMY\_HPP13, polypeptide determined that it comprised the following features: one transmembrane domain (TM1) located from about amino acid 225 to about amino acid 243 (TM1; SEQ ID NO:38) of SEQ ID NO:2 (Figures 1A-B) represented by double underlining.

Figure 2 shows the partial polynucleotide sequence (SEQ ID NO:3) and partial deduced amino acid sequence (SEQ ID NO:4) of the novel human phosphatase, BMY\_HPP13, of the present invention. The standard one-letter abbreviation for amino acids is used to illustrate the deduced amino acid sequence. The polynucleotide

sequence of BMY\_HPP13 contains a sequence of 624 nucleotides (SEQ ID NO:3), encoding a polypeptide of 208 amino acids (SEQ ID NO:4).

Figure 3A shows the regions of identity between the encoded full-length human phosphatase protein BMY\_HPP13 (SEQ ID NO:2), to the human CDC25B phosphatase protein (pdb1qb0.A.-; Genbank Accession No:gi|NM\_004358; SEQ ID NO:7). The alignment was performed using the FASTA algorithm (Pearson, *et. al.* 1990).

Figure 3B shows the regions of identity between the encoded full-length human phosphatase protein BMY\_HPP13 (SEQ ID NO:2), to the human tyrosine phosphatase Shp-2 protein (Target; Genbank Accession No:gi|4558224; SEQ ID NO:16). The alignment was performed using the FASTA algorithm (Pearson, *et. al.* 1990).

Figure 4 show an alignment of the BMY\_HPP13 polypeptide of the present invention (SEQ ID NO:2) with the corresponding genomic sequence (Genbank Accession No. AC06831; SEQ ID NO:12). The alignment was performed using the Genewisedb algorithm using default parameters (Genome Res. 10:547-8 (2000)). The alignment illustrates the predicted locations of each of the introns within the genomic sequence, and how the intron location relates to the BMY\_HPP13 polypeptide. As shown, the Genewise algorithm predicts the presence of two introns beginning at nucleotide 1369 to nucleotide 1970, and beginning at nucleotide 2065 to nucleotide 2090 of the AC06831 genomic sequence (“intron 1” and “intron 2”; respectively).

Figure 5 shows an expanded expression profile of the human dual specificity phosphatase, BMY\_HPP13. The figure illustrates the relative expression level of BMY\_HPP13 amongst various mRNA tissue sources. As shown, the BMY\_HPP13 polypeptide was expressed significantly in a majority of the tissues tested. Expression data was obtained by measuring the steady state BMY\_HPP13 mRNA levels by quantitative PCR using the PCR primer pair provided as SEQ ID NO:13 and 14, and Taqman probe (SEQ ID NO:15) as described in Example 4 herein.

Figure 6 shows a table illustrating the percent identity and percent similarity between the BMY\_HPP13 (SEQ ID NO:2), and the CDC25B phosphatase protein (pdb1qb0.A.-; Genbank Accession No:gi|NM\_004358; SEQ ID NO:7). The percent identity and percent similarity values were determined based upon the GAP algorithm

(GCG suite of programs; and Henikoff, S. and Henikoff, J. G., Proc. Natl. Acad. Sci. USA 89: 10915-10919(1992)) using the following parameters: gap weight = 8, and length weight = 2.

Figure 7 shows a sequence alignment of the translated sequence of the  
 5 BMY\_HPP13 polypeptide of the present invention (SEQ ID NO:2) with *human*  
 protein-tyrosine phosphatase 1B (Protein Data Bank entry 1AAX; Genbank  
 Accession No. gi|2981942; SEQ ID NO:40; Y.A.Puius et al., Proc.Nat.Acad.Sci.  
 USA, 94: 13420 (1997)). The alignment was used as the basis for building the  
 BMY\_HPP13 homology model described herein. The coordinates of the  
 10 BMY\_HPP13 model are provided in Table IV. Amino acids that are predicted to  
 comprise the putative binding site of BMY\_HPP13 are highlighted with an asterisk  
 (\*) below the BMY\_HPP13 sequence. Amino acids that were determined to comprise  
 the binding site of protein-tyrosine phosphatase 1B are highlighted with a plus  
 (+) sign above the 1AAX sequence. As shown, the majority of residues essential for  
 15 ligand binding are conserved between 1AAX and BMY\_HPP13.

Figure 8 shows the three-dimensional homology model of the BMY\_HPP13  
 polypeptide of the present invention (SEQ ID NO:2). The model is based upon an  
 alignment to a structural homologue *human* protein-tyrosine phosphatase 1B (Protein  
 Data Bank entry 1AAX; Genbank Accession No. gi|2981942; SEQ ID NO:40;  
 20 Y.A.Puius et al., Proc.Nat.Acad.Sci. USA, 94: 13420 (1997)) that was used as the  
 basis for building the BMY\_HPP13 homology model. The active site side chains that  
 are conserved or are homologous to those in the template PTP1B, 1AAX, are  
 highlighted. The coordinates of the BMY\_HPP13 model are provided in Table IV.

Table I provides a summary of the novel polypeptides and their encoding  
 25 polynucleotides of the present invention.

Table II illustrates the preferred hybridization conditions for the  
 polynucleotides of the present invention. Other hybridization conditions may be  
 known in the art or are described elsewhere herein.

Table III provides a summary of various conservative substitutions  
 30 encompassed by the present invention.

Table IV provides the structural coordinates of the three dimensional structure  
 of the BMY\_HPP13 polypeptide of the present invention (SEQ ID NO:2).

## DETAILED DESCRIPTION OF THE INVENTION

The present invention may be understood more readily by reference to the following detailed description of the preferred embodiments of the invention and the  
5 Examples included herein.

The invention provides a human polynucleotide sequence encoding a novel human phosphatase with substantial homology to the class of phosphatases known as phosphotyrosine or dual-specificity (P-Tyr, P-Ser and P-Thr) phosphatases. Members of this class of phosphatases have been implicated in a number of diseases and/or  
10 disorders, which include, but are not limited to, bone disorders, (Yoon, HK., Baylink, DJ., Lau, KH, Am. J. Nephrol., 20(2):153-62, ( 2000)), disease resistance to pathogens, reproductive disorders (Gloria, Bottini, F., Nicotra, M., Lucarini, N., Borgiani, P., La, Torre, M., Amante, A., Gimelfarb, A., Bottini, E, Dis. Markers., 12(4):261-9, (1996)), neural disorders (Shimohama, S., Fujimoto, S., Taniguchi, T.,  
15 Kameyama, M., Kimura, J. Ann, Neurol., 33(6):616-21, (1993)), prostate cancer (Nguyen, L., Chapdelaine, A., and Chevalier, S., Clin. Chem. 36(8 Pt 1): 1450-5 (1990)), immune disorders, particularly those relating to haematopoietic cell development, apoptosis, activation, and nonresponsiveness (Frearson, JA., Alexander, DR, Bioessays., 19(5): 417-27 (1997)), etc.

20 In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of  
25 matter, or particular cell is not the original environment of the polynucleotide. The term "isolated" does not refer to genomic or cDNA libraries, whole cell total or mRNA preparations, genomic DNA preparations (including those separated by electrophoresis and transferred onto blots), sheared whole cell genomic DNA preparations or other compositions where the art demonstrates no distinguishing  
30 features of the polynucleotide/sequences of the present invention.

In specific embodiments, the polynucleotides of the invention are at least 15, at least 30, at least 50, at least 100, at least 125, at least 500, or at least 1000

continuous nucleotides but are less than or equal to 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, 7.5 kb, 5 kb, 2.5 kb, 2.0 kb, or 1 kb, in length. In a further embodiment, polynucleotides of the invention comprise a portion of the coding sequences, as disclosed herein, but do not comprise all or a portion of any intron. In another  
5 embodiment, the polynucleotides comprising coding sequences do not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene of interest in the genome). In other embodiments, the polynucleotides of the invention do not contain the coding sequence of more than 1000, 500, 250, 100, 50, 25, 20, 15, 10, 5, 4, 3, 2, or 1 genomic flanking gene(s).

10 As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:2 or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without a signal sequence, the secreted protein  
15 coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

In the present invention, the full length sequence identified as SEQ ID NO:1  
20 was often generated by overlapping sequences contained in one or more clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:1 was deposited with the American Type Culture Collection ("ATCC"). As shown in Table I, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard,  
25 Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure. The deposited clone is inserted in the pSport plasmid (Life Technologies) using SalI and NotI restriction sites as described herein.

30 Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer (such as the Model 373, preferably a Model 3700, from Applied

Biosystems, Inc.), and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein  
5 may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known  
10 in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an  
15 insertion or deletion.

Using the information provided herein, such as the nucleotide sequence provided as SEQ ID NO: 1, a nucleic acid molecule of the present invention encoding a human phosphatase polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting  
20 material.

A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an  
25 overnight incubation at 42 degree C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65 degree C.

Also contemplated are nucleic acid molecules that hybridize to the  
30 polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower

percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37 degree C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M NaH<sub>2</sub>PO<sub>4</sub>; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50 degree C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone generated using oligo dT as a primer).

The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for

example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

The polypeptide of the present invention can be composed of amino acids  
5 joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more  
10 detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of  
15 modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a  
20 heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation,  
25 iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, *Proteins - Structure and Molecular Properties*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); *Posttranslational Covalent Modification of Proteins*, B. C. Johnson, Ed., Academic Press, New York,  
30 pgs. 1-12 (1983); Seifter et al., *Meth Enzymol* 182:626-646 (1990); Rattan et al., *Ann NY Acad Sci* 663:48-62 (1992).)

As will be appreciated by the skilled practitioner, should the amino acid fragment comprise an antigenic epitope, for example, biological function *per se* need not be maintained. The terms BMY\_HPP13 polypeptide and BMY\_HPP13 protein are used interchangeably herein to refer to the encoded product of the BMY\_HPP13  
5 nucleic acid sequence according to the present invention.

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences are identified by an integer specified in Table I.

"A polypeptide having biological activity" refers to polypeptides exhibiting  
10 activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the  
15 present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention).

It is another aspect of the present invention to provide modulators of the  
20 BMY\_HPP13 protein and BMY\_HPP13 peptide targets which can affect the function or activity of BMY\_HPP13 in a cell in which BMY\_HPP13 function or activity is to be modulated or affected. In addition, modulators of BMY\_HPP13 can affect downstream systems and molecules that are regulated by, or which interact with, BMY\_HPP13 in the cell. Modulators of BMY\_HPP13 include compounds, materials,  
25 agents, drugs, and the like, that antagonize, inhibit, reduce, block, suppress, diminish, decrease, or eliminate BMY\_HPP13 function and/or activity. Such compounds, materials, agents, drugs and the like can be collectively termed "antagonists". Alternatively, modulators of BMY\_HPP13 include compounds, materials, agents, drugs, and the like, that agonize, enhance, increase, augment, or amplify  
30 BMY\_HPP13 function in a cell. Such compounds, materials, agents, drugs and the like can be collectively termed "agonists".

As used herein the terms “modulate” or “modulates” refer to an increase or decrease in the amount, quality or effect of a particular activity, DNA, RNA, or protein. The definition of “modulate” or “modulates” as used herein is meant to encompass agonists and/or antagonists of a particular activity, DNA, RNA, or protein.

5       The term “organism” as referred to herein is meant to encompass any organism referenced herein, though preferably to eukaryotic organisms, more preferably to mammals, and most preferably to humans.

10       The present invention encompasses the identification of proteins, nucleic acids, or other molecules, that bind to polypeptides and polynucleotides of the present invention (for example, in a receptor-ligand interaction). The polynucleotides of the present invention can also be used in interaction trap assays (such as, for example, that described by Ozenberger and Young (Mol Endocrinol., 9(10):1321-9, (1995); and Ann. N. Y. Acad. Sci., 7;766:279-81, (1995)).

15       The polynucleotide and polypeptides of the present invention are useful as probes for the identification and isolation of full-length cDNAs and/or genomic DNA which correspond to the polynucleotides of the present invention, as probes to hybridize and discover novel, related DNA sequences, as probes for positional cloning of this or a related sequence, as probe to “subtract-out” known sequences in the process of discovering other novel polynucleotides, as probes to quantify gene expression, and as probes for microarrays.

20       In addition, polynucleotides and polypeptides of the present invention may comprise one, two, three, four, five, six, seven, eight, or more membrane domains.

25       Also, in preferred embodiments the present invention provides methods for further refining the biological function of the polynucleotides and/or polypeptides of the present invention.

30       Specifically, the invention provides methods for using the polynucleotides and polypeptides of the invention to identify orthologs, homologs, paralog, variants, and/or allelic variants of the invention. Also provided are methods of using the polynucleotides and polypeptides of the invention to identify the entire coding region of the invention, non-coding regions of the invention, regulatory sequences of the invention, and secreted, mature, pro-, prepro-, forms of the invention (as applicable).

In preferred embodiments, the invention provides methods for identifying the glycosylation sites inherent in the polynucleotides and polypeptides of the invention, and the subsequent alteration, deletion, and/or addition of said sites for a number of desirable characteristics which include, but are not limited to, augmentation of protein folding, inhibition of protein aggregation, regulation of intracellular trafficking to organelles, increasing resistance to proteolysis, modulation of protein antigenicity, and mediation of intercellular adhesion.

In further preferred embodiments, methods are provided for evolving the polynucleotides and polypeptides of the present invention using molecular evolution techniques in an effort to create and identify novel variants with desired structural, functional, and/or physical characteristics.

The present invention further provides for other experimental methods and procedures currently available to derive functional assignments. These procedures include but are not limited to spotting of clones on arrays, micro-array technology, PCR based methods (e.g., quantitative PCR), anti-sense methodology, gene knockout experiments, and other procedures that could use sequence information from clones to build a primer or a hybrid partner.

As used herein the terms "modulate or modulates" refer to an increase or decrease in the amount, quality or effect of a particular activity, DNA, RNA, or protein.

#### Polynucleotides and Polypeptides of the Invention

##### Features of the Polypeptide Encoded by Gene No:1

The polypeptide corresponding to this gene provided as SEQ ID NO:2 (Figures 1A-B), encoded by the polynucleotide sequence according to SEQ ID NO:1 (Figures 1A-B), and/or encoded by the polynucleotide contained within the deposited clone, BMY\_HPP13, has significant homology at the nucleotide and amino acid level to a number of phosphatases, which include, for example, the human CDC25B protein (pdb1qb0.A.-; Genbank Accession No:gi|NM\_004358; SEQ ID NO:7); and the human Shp-2 protein (Genbank Accession No:gi|4558224; SEQ ID NO:16); as determined by CLUSTALW. An alignment of the human phosphatase polypeptide with these proteins is provided in Figures 3A-B. The conserved catalytic residues are noted.

The BMY\_HPP13 polypeptide was determined to share 18.1% identity and 22.7% similarity with the human CDC25B protein (pdb1qb0.A.-; Genbank Accession No:gi|NM\_004358; SEQ ID NO:7); and 30.0% identity and 50.0% similarity with the human Shp-2 protein (Genbank Accession No:gi|4558224; SEQ ID NO:16).

5           The human CDC25B protein (pdb1qb0.A.-; Genbank Accession No:gi|NM\_004358; SEQ ID NO:7) is a member of the CDC25 family of phosphatases and has been determined to activate the cyclin dependent kinase CDC2 by removing two phosphate groups. CDC25B is required for entry into mitosis. CDC25B shuttles between the nucleus and the cytoplasm due to nuclear localization and nuclear export  
10       signals. The protein is nuclear in the M and G1 phases of the cell cycle and moves to the cytoplasm during S and G2. CDC25B has oncogenic properties, although its role in tumor formation has not been determined. Additional information relative to CDC25B may be obtained by reference to the following, non-limiting publications (New Biol. 3 (10), 959-968 (1991); Cell 67 (6), 1181-1194 (1991); Genomics 15 (3),  
15       693-694 (1993); Genomics 18 (1), 144-147 (1993); Science 269 (5230), 1575-1577 (1995); Oncogene 14 (20), 2485-2495 (1997); Biochem. Biophys. Res. Commun. 260 (2), 510-515 (1999); J. Cell Biol. 146 (3), 573-584 (1999); J. Mol. Biol. 293 (3), 559-568 (1999); and/or Oncogene 19 (18), 2179-2185 (2000); which are hereby incorporated herein by reference in their entirety.

20           Additional information relative to the structure of the human Shp-2 protein (Genbank Accession No:gi|4558224; SEQ ID NO:16) may be found by reference to the following publication: Cell 1998 Feb 20;92(4):441-50; which is hereby incorporated herein by reference in its entirety.

25           Analysis of the BMY\_HPP13 phosphatase polypeptide sequence led to the identification of a putative transmembrane domain located from about amino acid 225 to about amino acid 243 of SEQ ID NO:2. In this context, the term "about" may be construed to mean 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids beyond the N-Terminus and/or C-terminus of the above referenced transmembrane domain polypeptides.

30           In preferred embodiments, the following transmembrane domain polypeptide is encompassed by the present invention: PLNICVFILLLVFIVVKCF (SEQ ID NO:38). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of these BMY\_HPP13 transmembrane domain

polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

The transmembrane domain of BMY\_HPP13 is thought to anchor the polypeptide to the membrane such that the N-terminus of the polypeptide is on the outside of the cell. The polypeptide corresponding to amino acids from about 1 to about 224 of SEQ ID NO:2, and fragments thereof, are encompassed by the present invention. The present invention also encompasses the use of these BMY\_HPP13 polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

Protein threading and molecular modeling of BMY\_HPP13 suggest that BMY\_HPP13 has a three dimensional fold similar to that of the *human* protein-tyrosine phosphatase 1B (Protein Data Bank entry 1AAX; Genbank Accession No. gi|2981942; SEQ ID NO:40; Y.A.Puius et al., Proc.Nat.Acad.Sci. USA, 94: 13420 (1997)). The three dimensional structure of the human BMY\_HPP13 phosphatase polypeptide of the present invention is provided in Figure 8 An alignment of the BMY\_HPP13 polypeptide sequence to the *human* protein-tyrosine phosphatase 1B polypeptide is shown in Figure 7. The conserved ligand binding domain amino acids are noted in Figure 7.

The three dimensional crystallographic structure for numerous protein-tyrosine phosphatases (PTPases) reported and are deposited into the Protein Data Bank (Bernstein et. al., 1977 & Berman et. al., 2000).

The protein-tyrosine phosphatase 1B (PTP1B) structure is a structural prototype for the protein-tyrosine phosphatase family. PTP1B is a prototypical intracellular protein-tyrosine phosphatase and is found in a wide variety of human tissues. The structure of PTP1B (Puius et al., 1997)) was obtained from the Protein Data Bank (PDB) and has the PDB code 1AAX. The structure is representative for this class of enzymes E.C. 3.1.3.48. The structure contains parallel and anti-parallel beta strands composing the central beta sheet. Alpha helices surround the core sheet and the three critical loops that compose the binding site give the individual phosphatases their selectivity. The signature loop binds to the phosphate group and contains a catalytic cysteine. The WpD loop contains an aspartate residue used as the general acid/base during catalysis and defines one boundary of the binding pocket.

The WpD loop also contains residues that interact with substrate proximal to the phosphorylated tyrosine. The phosphatase active site is located within a cleft that is from 6-9 angstroms deep. Aryl side chains line the cleft and sandwich the phosphorylated substrate (e.g. phosphorylated tyrosine; pTyr). The third important  
 5 loop, the phosphate-binding loop forms the floor of the active site.

The sequence alignment (Figure 7) used as a template for creating the three-dimensional model of HPP\_BMY\_13 protein phosphatase domain has 22% sequence identity between the catalytic domain of HPP\_BMY\_13 and human PTP1B, PDB code 1AAX. For the protein-tyrosine phosphatase family of intracellular protein  
 10 phosphatases, the functionally important residues are located on three loops. The signature loop binds to the phosphate group and contains catalytic residues. The WpD loop contains Trp-179 and Asp-181 that provide the general acid/base during catalysis. In addition this loop defines the extent of the binding pocket. The WpD loop also contains residues that interact with substrate proximal to the phosphorylated  
 15 tyrosine. Like other members of this family, the phosphatase active site is located within a cleft that is from 6-9 angstroms deep. Aryl side chains Tyr-46 and F-182 line the cleft and sandwich the phosphorylated substrate (e.g. phosphorylated tyrosine; pTyr). The third loop, the phosphate-binding loop forms the floor of the active site. These residues are highlighted in the sequence alignment provided in Figure 7. The  
 20 other active site residues are also highlighted in Figure 7 and it is clear that several of the active site residues are completely conserved. Figure 8 shows the structure of the HPP\_BMY\_13 and has highlighted the active site side chains that are conserved or are homologous to those in the template PTP1B, 1AAX.

Homology models are useful when there is no experimental information  
 25 available on the protein of interest. A three dimensional model can be constructed on the basis of the known structure of a homologous protein (Greer *et. al.*, 1991, Lesk, *et. al.*, 1992, Levitt, 1992, Cardozo, *et. al.*, 1995, Sali, *et. al.*, 1995).

Those of skill in the art will understand that a homology model is constructed on the basis of first identifying a template, or, protein of known structure which is  
 30 similar to the protein without known structure. This can be accomplished by through pairwise alignment of sequences using such programs as FASTA (Pearson, *et. al.* 1990) and BLAST (Altschul, *et. al.*, 1990). In cases where sequence similarity is high

(greater than 30 %) these pairwise comparison methods may be adequate. Likewise, multiple sequence alignments or profile-based methods can be used to align a query sequence to an alignment of multiple (structurally and biochemically) related proteins. When the sequence similarity is low, more advanced techniques are used such as fold recognition (protein threading; Hendlich, *et. al.*, 1990, Koppensteiner *et. al.* 2000, Sippl & Weitckus 1992, Sippl 1993), where the compatibility of a particular sequence with the three dimensional fold of a potential template protein is gauged on the basis of a knowledge-based potential. Following the initial sequence alignment, the query template can be optimally aligned by manual manipulation or by incorporation of other features (motifs, secondary structure predictions, and allowed sequence conservation). Next, structurally conserved regions can be identified and are used to construct the core secondary structure (Levitt, 1992, Sali, *et. al.*, 1995) elements in the three dimensional model. Variable regions, called “unconserved regions” and loops can be added using knowledge-based techniques. The complete model with variable regions and loops can be refined performing forcefield calculations (Sali, *et. al.*, 1995, Cardozo, *et. al.*, 1995).

For BMV\_HPP13 a pairwise alignment generated by FASTA was used to align the sequence of BMV\_HPP13 with the sequence of the *human* protein-tyrosine phosphatase 1B (Protein Data Bank entry 1AAX; Genbank Accession No. gi|2981942; SEQ ID NO:40; Y.A.Puius *et al.*, Proc.Nat.Acad.Sci. USA, 94: 13420 (1997)). The alignment of BMV\_HPP13 with PDB entry 1AAX is set forth in Figure 7. In this invention, the homology model of BMV\_HPP13 was derived from the sequence alignment set forth in Figure 7. An overall atomic model including plausible sidechain orientations was generated using the program LOOK (Levitt, 1992). The three dimensional model for BMV\_HPP13 is defined by the set of structure coordinates as set forth in Table IV and is shown in Figure 8 rendered by backbone secondary structures.

The term “structure coordinates” refers to Cartesian coordinates generated from the building of a homology model.

Those of skill in the art will understand that a set of structure coordinates for a protein is a relative set of points that define a shape in three dimensions. Thus, it is possible that an entirely different set of coordinates could define a similar or identical

shape. Moreover, slight variations in the individual coordinates, as emanate from generation of similar homology models using different alignment templates (i.e., other than the structure coordinates of 1AAX), and/or using different methods in generating the homology model, will have minor effects on the overall shape. Variations in coordinates may also be generated because of mathematical manipulations of the structure coordinates. For example, the structure coordinates set forth in Table IV could be manipulated by fractionalization of the structure coordinates; integer additions or subtractions to sets of the structure coordinates, inversion of the structure coordinates or any combination of the above.

Various computational analyses are therefore necessary to determine whether a molecule or a portion thereof is sufficiently similar to all or parts of BMY\_HPP13 described above as to be considered the same. Such analyses may be carried out in current software applications, such as INSIGHTII (Accelrys Inc., San Diego, CA) version 2000 as described in the User's Guide, online ([www.accelrys.com](http://www.accelrys.com)) or software applications available in the SYBYL software suite (Tripos Inc., St. Louis, MO).

Using the superimposition tool in the program INSIGHTII comparisons can be made between different structures and different conformations of the same structure. The procedure used in INSIGHTII to compare structures is divided into four steps: 1) load the structures to be compared; 2) define the atom equivalencies in these structures; 3) perform a fitting operation; and 4) analyze the results. Each structure is identified by a name. One structure is identified as the target (i.e., the fixed structure); the second structure (i.e., moving structure) is identified as the source structure. Since atom equivalency within INSIGHTII is defined by user input, for the purpose of this invention we will define equivalent atoms as protein backbone atoms (N, C $\alpha$ , C and O) for all conserved residues between the two structures being compared. We will also consider only rigid fitting operations. When a rigid fitting method is used, the working structure is translated and rotated to obtain an optimum fit with the target structure. The fitting operation uses an algorithm that computes the optimum translation and rotation to be applied to the moving structure, such that the root mean square difference of the fit over the specified pairs of equivalent atom is an absolute minimum. This number, given in angstroms, is reported by INSIGHTII.

For the purpose of this invention, any homology model of a BMY\_HPP13 that has a root mean square deviation of conserved residue backbone atoms (N, C $\alpha$ , C, O) of less than 3.0 Å when superimposed on the relevant backbone atoms described by structure coordinates listed in Table IV are considered identical. More preferably, the  
5 root mean square deviation is less than about 2.0, 1.5, 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, or 0.1 Å.

The term "root mean square deviation" means the square root of the arithmetic mean of the squares of the deviations from the mean. It is a way to express the deviation or variation from a trend or object. For purposes of this invention, the "root  
10 mean square deviation" defines the variation in the backbone of a protein from the relevant portion of the backbone of BMY\_HPP13 as defined by the structure coordinates described herein.

This invention as embodied by the three-dimensional model enables the structure-based design of modulators of the biological function of BMY\_HPP13, as  
15 well as mutants with altered biological function and/or specificity.

The conservation of the amino acids in both of these functional sites and the overall 22% sequence identity emphasize the significance of the three-dimensional model. The conserved residues are located within the ligand binding domain. These functional site residues play critical roles in the mechanism of catalysis, substrate  
20 specificity and ligand binding.

The structure coordinates of a BMY\_HPP13 homology model portion thereof are stored in a machine-readable storage medium. Such data may be used for a variety of purposes, such as drug discovery and target prioritization and validation.

Accordingly, in one embodiment of this invention is provided a machine-  
25 readable data storage medium comprising a data storage material encoded with the structure coordinates set forth in Table IV.

For the first time, the present invention permits the use, through homology modeling based upon the sequence of BMY\_HPP13 (Figures 1A-B) of structure-based or rational drug design techniques to design, select, and synthesizes chemical  
30 entities that are capable of modulating the biological function of BMY\_HPP13. Comparison of the BMY\_HPP13 homology model with the structures of other the phosphatases, particularly dual specificity phosphatases, enables the use of rational or

structure based drug design methods to design, select or synthesize specific chemical modulators of BMY\_HPP13.

The three-dimensional model structure of the BMY\_HPP13 also provides methods for identifying modulators of biological function. Various methods or  
5 combination thereof can be used to identify these compounds.

Structure coordinates of the ligand binding domain defined above can also be used to identify structural and chemical features. Identified structural or chemical features can then be employed to design or select compounds as potential BMY\_HPP13 modulators. By structural and chemical features it is meant to include,  
10 but is not limited to, van der Waals interactions, hydrogen bonding interactions, charge interaction, hydrophobic interactions, and dipole interaction. Alternatively, or in conjunction, the three-dimensional structural model can be employed to design or select compounds as potential BMY\_HPP13 modulators. Compounds identified as potential BMY\_HPP13 modulators can then be synthesized and screened in an assay  
15 characterized by binding of a test compound to the BMY\_HPP13, or in characterizing BMY\_HPP13 deactivation in the presence of a small molecule. Examples of assays useful in screening of potential BMY\_HPP13 modulators include, but are not limited to, screening *in silico*, *in vitro* assays and high throughput assays. Finally, these methods may also involve modifying or replacing one or more amino acids from  
20 BMY\_HPP13 according to Table IV.

However, as will be understood by those of skill in the art upon this disclosure, other structure based design methods can be used. Various computational structure based design methods have been disclosed in the art.

For example, a number of computer modeling systems are available in which  
25 the sequence of the BMY\_HPP13 and the BMY\_HPP13 structure (i.e., atomic coordinates of BMY\_HPP13 and/or the atomic coordinates of the active site region as provided in Table IV) can be input. The computer system then generates the structural details of one or more these regions in which a potential BMY\_HPP13 modulator binds so that complementary structural details of the potential modulators can be  
30 determined. Design in these modeling systems is generally based upon the compound being capable of physically and structurally associating with BMY\_HPP13. In addition, the compound must be able to assume a conformation that allows it to

associate with BMY\_HPP13. Some modeling systems estimate the potential inhibitory or binding effect of a potential BMY\_HPP13 modulator prior to actual synthesis and testing.

5       Methods for screening chemical entities or fragments for their ability to  
associate with a given protein target are well known. Often these methods begin by  
visual inspection of the binding site on the computer screen. Selected fragments or  
chemical entities are then positioned in one or more positions and orientations within  
the active site region in BMY\_HPP13. Molecular docking is accomplished using  
software such as INSIGHTII, ICM (Molsoft LLC, La Jolla, CA), and SYBYL,  
10       following by energy minimization and molecular dynamics with standard molecular  
mechanic forcefields such as CHARMM and MMFF. Examples of computer  
programs which assist in the selection of chemical fragment or chemical entities  
useful in the present invention include, but are not limited to, GRID (Goodford,  
1985), AUTODOCK (Goodsell, 1990), and DOCK (Kuntz *et. al.* 1982).

15       Alternatively, compounds may be designed de novo using either an empty  
active site, ligand binding domain, or optionally including some portion of a known  
inhibitor. Methods of this type of design include, but are not limited to LUDI (Bohm  
1992), LeapFrog (Tripos Associates, St. Louis MO) and DOCK (Kuntz *et. al.*, 1982).  
Programs such as DOCK (Kuntz *et. al.* 1982) can be used with the atomic coordinates  
20       from the homology model to identify potential ligands from databases or virtual  
databases which potentially bind the in the active site region, and which may therefore  
be suitable candidates for synthesis and testing. The computer programs may utilize a  
combination of the following steps: a.) Selection of fragments or chemical entities  
from a database and then positioning the chemical entity in one or more orientations  
25       within the BMY\_HPP13 catalytic domain defined by Table IV; b.) characterization of  
the structural and chemical features of the chemical entity and active site including  
van der Waals interactions, hydrogen bonding interactions, charge interaction,  
hydrophobic bonding interaction, and dipole interactions; c.) Search databases for  
molecular fragments which can be joined to or replace the docked chemical entity and  
30       spatially fit into regions defined by the said BMY\_HPP13 catalytic domain or  
catalytic domain functional sites; and/or d.) Evaluate the docked chemical entity and  
fragments using a combination of scoring schemes which account for van der Waals

interactions, hydrogen bonding interactions, charge interaction, hydrophobic interactions

Databases that may be used include ACD (Molecular Designs Limited), Aldrich (Aldrich Chemical Company), NCI (National Cancer Institute),  
5 Maybridge(Maybridge Chemical Company Ltd), CCDC (Cambridge Crystallographic Data Center), CAST (Chemical Abstract Service), Derwent (Derwent Information Limited).

Upon selection of preferred chemical entities or fragments, their relationship to each other and BMY\_HPP13 can be visualized and then assembled into a single  
10 potential modulator. Programs useful in assembling the individual chemical entities include, but are not limited to SYBYL and LeapFrog (Tripos Associates, St. Louis MO), LUDI (Bohm 1992) as well as 3D Database systems (Martin 1992).

Additionally, the three-dimensional homology model of BMY\_HPP13 will aid in the design of mutants with altered biological activity. Site directed mutagenesis can  
15 be used to generate proteins with similar or varying degrees of biological activity compared to native BMY\_HPP13. This invention also relates to the generation of mutants or homologs of BMY\_HPP13. It is clear that molecular modeling using the three dimensional structure coordinates set forth in Table IV and visualization of the BMY\_HPP13 model, Figure 8 can be utilized to design homologs or mutant  
20 polypeptides of BMY\_HPP13 that have similar or altered biological activities, function or reactivities.

Based upon the strong homology and structural conservation to members of the phosphatase proteins, the polypeptide encoded by the human BMY\_HPP13 phosphatase of the present invention is expected to share at least some biological  
25 activity with phosphatase proteins, preferably with members of the novel phosphotyrosine/dual-specificity (P-Tyr, P-Ser and P-Thr) phosphatases, particularly the novel phosphotyrosine/dual-specificity (P-Tyr, P-Ser and P-Thr) phosphatases referenced herein.

The present invention encompasses the use of BMY\_HPP13 inhibitors and/or  
30 activators of BMY\_HPP13 activity for the treatment, detection, amelioration, or prevention of phosphatase associated disorders, including but not limited to metabolic diseases such as diabetes, in addition to neural and/or cardiovascular diseases and

disorders. The present invention also encompasses the use of BMY\_HPP13 inhibitors and/or activators of BMY\_HPP13 activity as immunosuppressive agents, anti-inflammatory agents, and/or anti-tumor agents

5 The present invention encompasses the use of BMY\_HPP13 phosphatase inhibitors, including, antagonists such as antisense nucleic acids, in addition to other antagonists, as described herein, in a therapeutic regimen to diagnose, prognose, treat, ameliorate, and/or prevent diseases where a kinase activity is insufficient. One, non-limiting example of a disease which may occur due to insufficient kinase activity are certain types of diabetes, where one or more kinases involved in the insulin receptor  
10 signal pathway may have insufficient activity or insufficient expression, for example.

Moreover, the present invention encompasses the use of BMY\_HPP13 phosphatase activators, and/or the use of the BMY\_HPP13 phosphatase gene or protein in a gene therapy regimen, as described herein, for the diagnoses, prognoses, treatment, amelioration, and/or prevention of diseases and/or disorders where a kinase  
15 activity is overly high, such as a cancer where a kinase oncogene product has excessive activity or excessive expression.

The present invention also encompasses the use of catalytically inactive variants of BMY\_HPP13 proteins, including fragments thereof, such as a protein therapeutic, or the use of the encoding polynucleotide sequence or as gene therapy,  
20 for example, in the diagnoses, prognosis, treatment, amelioration, and/or prevention of diseases or disorders where phosphatase activity is overly high.

The present invention encompasses the use of antibodies directed against the BMY\_HPP13 polypeptides, including fragment and/or variants thereof, of the present invention in diagnostics, as a biomarkers, and/or as a therapeutic agents.

25 The present invention encompasses the use of an inactive, non-catalytic, mutant of the BMY\_HPP13 phosphatase as a substrate trapping mutant to bind cellular phosphoproteins or a library of phosphopeptides to identify substrates of the BMY\_HPP13 polypeptides.

The present invention encompasses the use of the BMY\_HPP13 polypeptides,  
30 to identify inhibitors or activators of the BMY\_HPP13 phosphatase activity using either in vitro or 'virtual' (in silico) screening methods.

One embodiment of the invention relates to a method for identifying a compound as an activator or inhibitor of the BMY\_HPP13 phosphatase comprising the steps of: i.) contacting a BMY\_HPP13 phosphatase inhibitor or activator labeled with an analytically detectable reagent with the BMY\_HPP13 phosphatase under  
5 conditions sufficient to form a complex with the inhibitor or activator; ii.) contacting said complex with a sample containing a compound to be identified; iii) and identifying the compound as an inhibitor or activator by detecting the ability of the test compound to alter the amount of labeled known BMY\_HPP13 phosphatase inhibitor or activator in the complex.

10 Another embodiment of the invention relates to a method for identifying a compound as an activator or inhibitor of a BMY\_HPP13 phosphatase comprising the steps of: i.) contacting the BMY\_HPP13 phosphatase with a compound to be identified; and ii.) and measuring the ability of the BMY\_HPP13 phosphatase to remove phosphate from a substrate.

15 The present invention also encompasses a method for identifying a ligand for the BMY\_HPP13 phosphatase comprising the steps of: i.) contacting the BMY\_HPP13 phosphatase with a series of compounds under conditions to permit binding; and ii.) detecting the presence of any ligand-bound protein.

20 Preferably, the above referenced methods comprise the BMY\_HPP13 phosphatase in a form selected from the group consisting of whole cells, cytosolic cell fractions, membrane cell fractions, purified or partially purified forms. The invention also relates to recombinantly expressed BMY\_HPP13 phosphatase in a purified, substantially purified, or unpurified state. The invention further relates to BMY\_HPP13 phosphatase fused or conjugated to a protein, peptide, or other  
25 molecule or compound known in the art, or referenced herein.

The present invention also encompasses pharmaceutical composition of the BMY\_HPP13 phosphatase polypeptide comprising a compound identified by above referenced methods and a pharmaceutically acceptable carrier.

30 Expression profiling designed to measure the steady state mRNA levels encoding the BMY\_HPP13 polypeptide showed significantly high expression levels in the majority of tissues tested (as shown in Figure 5). However, the BMY\_HPP13

polypeptide showed slightly increased levels of expression in fallopian tube, lymph gland, and the small intestine.

The strong homology to dual specificity phosphatases, combined with the predominate localized expression in lymph gland tissue suggests a potential utility for  
5    BMY\_HPP13 polynucleotides and polypeptides in treating, diagnosing, prognosing, and/or preventing immune diseases and/or disorders. Representative uses are described in the “Immune Activity”, “Chemotaxis”, and “Infectious Disease” sections below, and elsewhere herein. Briefly, the strong expression in immune tissue indicates a role in regulating the proliferation; survival; differentiation; and/or activation of  
10   hematopoietic cell lineages, including blood stem cells.

The BMY\_HPP13 polypeptide may also be useful as a preventative agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as  
15   T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma. The BMY\_HPP13 polypeptide may be useful for  
20   modulating cytokine production, antigen presentation, or other processes, such as for boosting immune responses, etc.

Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem  
25   cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissuemarkers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the  
30   protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

The strong homology to dual specificity phosphatase proteins, combined with the localized expression in small intestine suggests the BMY\_HPP13 polynucleotides and polypeptides may be useful in treating, diagnosing, prognosing, and/or preventing gastrointestinal diseases and/or disorders, which include, but are not limited to, ulcers, irritable bowel syndrome, inflammatory bowel disease, diarrhea, traveler's diarrhea, drug-related diarrhea, polyps, absorption disorders, constipation, diverticulitis, vascular disease of the intestines, intestinal obstruction, intestinal infections, ulcerative colitis, Shigellosis, cholera, Crohn's Disease, amebiasis, enteric fever, Whipple's Disease, peritonitis, intrabdominal abscesses, hereditary hemochromatosis, gastroenteritis, viral gastroenteritis, food poisoning, mesenteric ischemia, mesenteric infarction, in addition to, metabolic diseases and/or disorders.

Moreover, polynucleotides and polypeptides, including fragments and/or antagonists thereof, have uses which include, directly or indirectly, treating, preventing, diagnosing, and/or prognosing susceptibility to the following, non-limiting, gastrointestinal infections: Salmonella infection, E.coli infection, E.coli O157:H7 infection, Shiga Toxin-producing E.coli infection, Campylobacter infection (e.g., Campylobacter fetus, Campylobacter upsaliensis, Campylobacter hyointestinalis, Campylobacter lari, Campylobacter jejuni, Campylobacter concisus, Campylobacter mucosalis, Campylobacter sputorum, Campylobacter rectus, Campylobacter curvus, Campylobacter sputorum, etc.), Heliobacter infection (e.g., Heliobacter cinaedi, Heliobacter fennelliae, etc.) Yersinia enterocolitica infection, Vibrio sp. Infection (e.g., Vibrio mimicus, Vibrio parahaemolyticus, Vibrio fluvialis, Vibrio furnissii, Vibrio hollisae, Vibrio vulnificus, Vibrio alginolyticus, Vibrio metschnikovii, Vibrio damsela, Vibrio cincinnatiensis, etc.) Aeromonas infection (e.g., Aeromonas hydrophila, Aeromonas sobira, Aeromonas caviae, etc.), Plesiomonas shigellides infection, Giardia infection (e.g., Giardia lamblia, etc.), Cryptosporidium infection, Listeria infection, Entamoeba histolytica infection, Rotavirus infection, Norwalk virus infection, Clostridium difficile infection, Clostridium perfringens infection, Staphylococcus infection, Bacillus infection, in addition to any other gastrointestinal disease and/or disorder implicated by the causative agents listed above or elsewhere herein.

The strong homology to dual specificity phosphatases, combined with the predominate localized expression in fallopian tube tissue suggests a potential utility for BMY\_HPP13 polynucleotides and polypeptides in treating, diagnosing, prognosing, and/or preventing female reproductive diseases and/or disorders.

5       The BMY\_HPP13 polypeptide has been shown to comprise two glycosylation sites according to the Motif algorithm (Genetics Computer Group, Inc.). As discussed more specifically herein, protein glycosylation is thought to serve a variety of functions including: augmentation of protein folding, inhibition of protein aggregation, regulation of intracellular trafficking to organelles, increasing resistance  
10       to proteolysis, modulation of protein antigenicity, and mediation of intercellular adhesion.

Asparagine glycosylation sites have the following consensus pattern, N-{P}-[ST]-{P}, wherein N represents the glycosylation site. However, it is well known that that potential N-glycosylation sites are specific to the consensus sequence Asn-Xaa-  
15       Ser/Thr. However, the presence of the consensus tripeptide is not sufficient to conclude that an asparagine residue is glycosylated, due to the fact that the folding of the protein plays an important role in the regulation of N-glycosylation. It has been shown that the presence of proline between Asn and Ser/Thr will inhibit N-glycosylation; this has been confirmed by a recent statistical analysis of glycosylation  
20       sites, which also shows that about 50% of the sites that have a proline C-terminal to Ser/Thr are not glycosylated. Additional information relating to asparagine glycosylation may be found in reference to the following publications, which are hereby incorporated by reference herein: Marshall R.D., Annu. Rev. Biochem. 41:673-702(1972); Pless D.D., Lennarz W.J., Proc. Natl. Acad. Sci. U.S.A. 74:134-  
25       138(1977); Bause E., Biochem. J. 209:331-336(1983); Gavel Y., von Heijne G., Protein Eng. 3:433-442(1990); and Miletich J.P., Broze G.J. Jr., J. Biol. Chem. 265:11397-11404(1990).

In preferred embodiments, the following asparagine glycosylation site polypeptide is encompassed by the present invention: KKYYGNGTRKSPeM (SEQ  
30       ID NO:17), and/or ANQAKNQSAEAKEA (SEQ ID NO:18). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of these BMY\_HPP13 asparagine glycosylation site

polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

The BMY\_HPP13 polypeptides of the present invention were determined to comprise several phosphorylation sites based upon the Motif algorithm (Genetics Computer Group, Inc.). The phosphorylation of such sites may regulate some biological activity of the BMY\_HPP13 polypeptide. For example, phosphorylation at specific sites may be involved in regulating the proteins ability to associate or bind to other molecules (e.g., proteins, ligands, substrates, DNA, etc.).

The BMY\_HPP13 polypeptide was predicted to comprise three PKC phosphorylation sites using the Motif algorithm (Genetics Computer Group, Inc.). In vivo, protein kinase C exhibits a preference for the phosphorylation of serine or threonine residues. The PKC phosphorylation sites have the following consensus pattern: [ST]-x-[RK], where S or T represents the site of phosphorylation and 'x' an intervening amino acid residue. Additional information regarding PKC phosphorylation sites can be found in Woodget J.R., Gould K.L., Hunter T., Eur. J. Biochem. 161:177-184(1986), and Kishimoto A., Nishiyama K., Nakanishi H., Uratsuji Y., Nomura H., Takeyama Y., Nishizuka Y., J. Biol. Chem. 260:12492-12499(1985); which are hereby incorporated by reference herein.

In preferred embodiments, the following PKC phosphorylation site polypeptides are encompassed by the present invention: PRATWTLKLDGNL (SEQ ID NO:19), FSSDSTMRILSNL (SEQ ID NO:20), and/or YYGNGTRKSPEMP (SEQ ID NO:21). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of these BMY\_HPP13 PKC phosphorylation site polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

The BMY\_HPP13 polypeptide was predicted to comprise four casein kinase II phosphorylation sites using the Motif algorithm (Genetics Computer Group, Inc.). Casein kinase II (CK-2) is a protein serine/threonine kinase whose activity is independent of cyclic nucleotides and calcium. CK-2 phosphorylates many different proteins. The substrate specificity [1] of this enzyme can be summarized as follows: (1) Under comparable conditions Ser is favored over Thr.; (2) An acidic residue (either Asp or Glu) must be present three residues from the C-terminal of the

phosphate acceptor site; (3) Additional acidic residues in positions +1, +2, +4, and +5 increase the phosphorylation rate. Most physiological substrates have at least one acidic residue in these positions; (4) Asp is preferred to Glu as the provider of acidic determinants; and (5) A basic residue at the N-terminal of the acceptor site decreases the phosphorylation rate, while an acidic one will increase it.

A consensus pattern for casein kinase II phosphorylation site is as follows: [ST]-x(2)-[DE], wherein 'x' represents any amino acid, and S or T is the phosphorylation site.

Additional information specific to casein kinase II phosphorylation sites may be found in reference to the following publication: Pinna L.A., *Biochim. Biophys. Acta* 1054:267-284(1990); which is hereby incorporated herein in its entirety.

In preferred embodiments, the following casein kinase II phosphorylation site polypeptide is encompassed by the present invention: WTWEQTFQELIQEA (SEQ ID NO:22), QILCHTYWEHWTSQ (SEQ ID NO:23), QKCSWSQYEMPEFS (SEQ ID NO:24), and/or KEAKGSGYEKLGPS (SEQ ID NO:25). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of these casein kinase II phosphorylation site polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

The BMY\_HPP13 polypeptide was predicted to comprise two N-myristoylation sites using the Motif algorithm (Genetics Computer Group, Inc.). An appreciable number of eukaryotic proteins are acylated by the covalent addition of myristate (a C14-saturated fatty acid) to their N-terminal residue via an amide linkage. The sequence specificity of the enzyme responsible for this modification, myristoyl CoA:protein N-myristoyl transferase (NMT), has been derived from the sequence of known N-myristoylated proteins and from studies using synthetic peptides. The specificity seems to be the following: i.) The N-terminal residue must be glycine; ii.) In position 2, uncharged residues are allowed; iii.) Charged residues, proline and large hydrophobic residues are not allowed; iv.) In positions 3 and 4, most, if not all, residues are allowed; v.) In position 5, small uncharged residues are allowed (Ala, Ser, Thr, Cys, Asn and Gly). Serine is favored; and vi.) In position 6, proline is not allowed.

A consensus pattern for N-myristoylation is as follows: G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}, wherein 'x' represents any amino acid, and G is the N-myristoylation site.

Additional information specific to N-myristoylation sites may be found in reference to the following publication: Towler D.A., Gordon J.I., Adams S.P., Glaser L., *Annu. Rev. Biochem.* 57:69-99(1988); and Grand R.J.A., *Biochem. J.* 258:625-638(1989); which is hereby incorporated herein in its entirety.

In preferred embodiments, the following N-myristoylation site polypeptides are encompassed by the present invention: EVSLEGSHDTANCEAC (SEQ ID NO:26), and/or GICGQGLKSCMTKPSK (SEQ ID NO:27). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of these N-myristoylation site polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

In preferred embodiments, the following N-terminal BMY\_HPP13 deletion polypeptides are encompassed by the present invention: M1-E246, V2-E246, V3-E246, D4-E246, F5-E246, W6-E246, T7-E246, W8-E246, E9-E246, Q10-E246, T11-E246, F12-E246, Q13-E246, E14-E246, L15-E246, I16-E246, Q17-E246, E18-E246, A19-E246, K20-E246, P21-E246, R22-E246, A23-E246, T24-E246, W25-E246, T26-E246, L27-E246, K28-E246, L29-E246, D30-E246, G31-E246, N32-E246, L33-E246, Q34-E246, L35-E246, D36-E246, C37-E246, L38-E246, A39-E246, Q40-E246, G41-E246, W42-E246, K43-E246, Q44-E246, Y45-E246, Q46-E246, Q47-E246, R48-E246, A49-E246, F50-E246, G51-E246, W52-E246, F53-E246, R54-E246, C55-E246, S56-E246, S57-E246, C58-E246, Q59-E246, R60-E246, S61-E246, W62-E246, A63-E246, S64-E246, A65-E246, Q66-E246, V67-E246, Q68-E246, I69-E246, L70-E246, C71-E246, H72-E246, T73-E246, Y74-E246, W75-E246, E76-E246, H77-E246, W78-E246, T79-E246, S80-E246, Q81-E246, G82-E246, Q83-E246, V84-E246, R85-E246, M86-E246, R87-E246, L88-E246, F89-E246, G90-E246, Q91-E246, R92-E246, C93-E246, Q94-E246, K95-E246, C96-E246, S97-E246, W98-E246, S99-E246, Q100-E246, Y101-E246, E102-E246, M103-E246, P104-E246, E105-E246, F106-E246, S107-E246, S108-E246, D109-E246, S110-E246, T111-E246, M112-E246, R113-E246, I114-E246, L115-E246, S116-E246, N117-E246, L118-E246, V119-E246, Q120-E246, H121-E246, I122-E246, L123-

E246, K124-E246, K125-E246, Y126-E246, Y127-E246, G128-E246, N129-E246,  
 G130-E246, T131-E246, R132-E246, K133-E246, S134-E246, P135-E246, E136-  
 E246, M137-E246, P138-E246, V139-E246, I140-E246, L141-E246, E142-E246,  
 V143-E246, S144-E246, L145-E246, E146-E246, G147-E246, S148-E246, H149-  
 5 E246, D150-E246, T151-E246, A152-E246, N153-E246, C154-E246, E155-E246,  
 A156-E246, C157-E246, T158-E246, L159-E246, G160-E246, I161-E246, C162-  
 E246, G163-E246, Q164-E246, G165-E246, L166-E246, K167-E246, S168-E246,  
 C169-E246, M170-E246, T171-E246, K172-E246, P173-E246, S174-E246, K175-  
 E246, S176-E246, L177-E246, L178-E246, P179-E246, H180-E246, L181-E246,  
 10 K182-E246, T183-E246, G184-E246, N185-E246, S186-E246, S187-E246, P188-  
 E246, G189-E246, I190-E246, G191-E246, A192-E246, V193-E246, Y194-E246,  
 L195-E246, A196-E246, N197-E246, Q198-E246, A199-E246, K200-E246, N201-  
 E246, Q202-E246, S203-E246, A204-E246, E205-E246, A206-E246, K207-E246,  
 E208-E246, A209-E246, K210-E246, G211-E246, S212-E246, G213-E246, Y214-  
 15 E246, E215-E246, K216-E246, L217-E246, G218-E246, P219-E246, S220-E246,  
 R221-E246, D222-E246, P223-E246, D224-E246, P225-E246, L226-E246, N227-  
 E246, I228-E246, C229-E246, V230-E246, F231-E246, I232-E246, L233-E246,  
 L234-E246, L235-E246, V236-E246, F237-E246, I238-E246, V239-E246, and/or  
 V240-E246 of SEQ ID NO:2. Polynucleotide sequences encoding these polypeptides  
 20 are also provided. The present invention also encompasses the use of these N-terminal  
 BMY\_HPP13 deletion polypeptides as immunogenic and/or antigenic epitopes as  
 described elsewhere herein.

In preferred embodiments, the following C-terminal BMY\_HPP13 deletion  
 polypeptides are encompassed by the present invention: M1-E246, M1-S245, M1-  
 25 T244, M1-F243, M1-C242, M1-K241, M1-V240, M1-V239, M1-I238, M1-F237,  
 M1-V236, M1-L235, M1-L234, M1-L233, M1-I232, M1-F231, M1-V230, M1-C229,  
 M1-I228, M1-N227, M1-L226, M1-P225, M1-D224, M1-P223, M1-D222, M1-R221,  
 M1-S220, M1-P219, M1-G218, M1-L217, M1-K216, M1-E215, M1-Y214, M1-  
 G213, M1-S212, M1-G211, M1-K210, M1-A209, M1-E208, M1-K207, M1-A206,  
 30 M1-E205, M1-A204, M1-S203, M1-Q202, M1-N201, M1-K200, M1-A199, M1-  
 Q198, M1-N197, M1-A196, M1-L195, M1-Y194, M1-V193, M1-A192, M1-G191,  
 M1-I190, M1-G189, M1-P188, M1-S187, M1-S186, M1-N185, M1-G184, M1-T183,

M1-K182, M1-L181, M1-H180, M1-P179, M1-L178, M1-L177, M1-S176, M1-K175, M1-S174, M1-P173, M1-K172, M1-T171, M1-M170, M1-C169, M1-S168, M1-K167, M1-L166, M1-G165, M1-Q164, M1-G163, M1-C162, M1-I161, M1-G160, M1-L159, M1-T158, M1-C157, M1-A156, M1-E155, M1-C154, M1-N153,  
 5 M1-A152, M1-T151, M1-D150, M1-H149, M1-S148, M1-G147, M1-E146, M1-L145, M1-S144, M1-V143, M1-E142, M1-L141, M1-I140, M1-V139, M1-P138, M1-M137, M1-E136, M1-P135, M1-S134, M1-K133, M1-R132, M1-T131, M1-G130, M1-N129, M1-G128, M1-Y127, M1-Y126, M1-K125, M1-K124, M1-L123, M1-I122, M1-H121, M1-Q120, M1-V119, M1-L118, M1-N117, M1-S116, M1-L115,  
 10 M1-I114, M1-R113, M1-M112, M1-T111, M1-S110, M1-D109, M1-S108, M1-S107, M1-F106, M1-E105, M1-P104, M1-M103, M1-E102, M1-Y101, M1-Q100, M1-S99, M1-W98, M1-S97, M1-C96, M1-K95, M1-Q94, M1-C93, M1-R92, M1-Q91, M1-G90, M1-F89, M1-L88, M1-R87, M1-M86, M1-R85, M1-V84, M1-Q83, M1-G82, M1-Q81, M1-S80, M1-T79, M1-W78, M1-H77, M1-E76, M1-W75, M1-Y74, M1-T73, M1-H72, M1-C71, M1-L70, M1-I69, M1-Q68, M1-V67, M1-Q66, M1-A65, M1-S64, M1-A63, M1-W62, M1-S61, M1-R60, M1-Q59, M1-C58, M1-S57, M1-S56, M1-C55, M1-R54, M1-F53, M1-W52, M1-G51, M1-F50, M1-A49, M1-R48, M1-Q47, M1-Q46, M1-Y45, M1-Q44, M1-K43, M1-W42, M1-G41, M1-Q40, M1-A39, M1-L38, M1-C37, M1-D36, M1-L35, M1-Q34, M1-L33, M1-N32, M1-G31,  
 15 M1-D30, M1-L29, M1-K28, M1-L27, M1-T26, M1-W25, M1-T24, M1-A23, M1-R22, M1-P21, M1-K20, M1-A19, M1-E18, M1-Q17, M1-I16, M1-L15, M1-E14, M1-Q13, M1-F12, M1-T11, M1-Q10, M1-E9, M1-W8, and/or M1-T7 of SEQ ID NO:2. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these C-terminal BMY\_HPP13 deletion  
 25 polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:1 and may have been publicly available prior to conception of  
 30 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention

are one or more polynucleotides consisting of a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 975 of SEQ ID NO:1, b is an integer between 15 to 989, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:1, and where b is greater than or equal to

5 a+14.

TABLE I

Gene No.	CDNA CloneID	ATCC Deposit No. Z and Date	Vector	NT SEQ ID. No. X	Total NT Seq of Clone	5' NT of Start Codon of ORF	3' NT of ORF	AA Seq ID No. Y	Total AA of ORF
1.	BMV_HPP 13	PTA-4803 11/14/02		1	989	26	763	2	246

Table I summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table I and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually several overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

10

15

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq. Of Clone" refers to the total number of nucleotides in the clone contig identified by "Gene No." The deposited clone may contain all or most of the sequence of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon of ORF."

20

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

25

The total number of amino acids within the open reading frame of SEQ ID NO:Y is identified as "Total AA of ORF".

SEQ ID NO:X (where X may be any of the polynucleotide sequences disclosed in the sequence listing) and the translated SEQ ID NO:Y (where Y may be any of the polypeptide sequences disclosed in the sequence listing) are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further herein. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used, for example, to generate antibodies which bind specifically to proteins containing the polypeptides and the proteins encoded by the cDNA clones identified in Table I.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides may cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:1 and the predicted translated amino acid sequence identified as SEQ ID NO:2, but also a sample of plasmid DNA containing a cDNA of the invention deposited with the ATCC, as set forth in Table I. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited cDNA, collecting the protein, and determining its sequence.

The present invention also relates to the genes corresponding to SEQ ID NO:1, SEQ ID NO:2, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and  
5 identifying or amplifying the corresponding gene from appropriate sources of genomic material.

Also provided in the present invention are species homologs, allelic variants, and/or orthologs. The skilled artisan could, using procedures well-known in the art, obtain the polynucleotide sequence corresponding to full-length genes (including, but  
10 not limited to the full-length coding region), allelic variants, splice variants, orthologs, and/or species homologues of genes corresponding to SEQ ID NO:1, SEQ ID NO:2, or a deposited clone, relying on the sequence from the sequences disclosed herein or the clones deposited with the ATCC. For example, allelic variants and/or species homologues may be isolated and identified by making suitable probes or primers  
15 which correspond to the 5', 3', or internal regions of the sequences provided herein and screening a suitable nucleic acid source for allelic variants and/or the desired homologue.

The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly  
20 produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the protein, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include  
25 an additional amino acid sequence which contains secretory or leader sequences, prosequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced  
30 version of a polypeptide, can be substantially purified using techniques described herein or otherwise known in the art, such as, for example, by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention

also can be purified from natural, synthetic or recombinant sources using protocols described herein or otherwise known in the art, such as, for example, antibodies of the invention raised against the full-length form of the protein.

5 The present invention provides a polynucleotide comprising, or alternatively consisting of, the sequence identified as SEQ ID NO:1, and/or a cDNA provided in ATCC Deposit No. Z:. The present invention also provides a polypeptide comprising, or alternatively consisting of, the sequence identified as SEQ ID NO:2, and/or a polypeptide encoded by the cDNA provided in ATCC Deposit NO:Z. The present invention also provides polynucleotides encoding a polypeptide comprising, or  
10 alternatively consisting of the polypeptide sequence of SEQ ID NO:2, and/or a polypeptide sequence encoded by the cDNA contained in ATCC Deposit No:Z.

Preferably, the present invention is directed to a polynucleotide comprising, or alternatively consisting of, the sequence identified as SEQ ID NO:1, and/or a cDNA provided in ATCC Deposit No.: that is less than, or equal to, a polynucleotide  
15 sequence that is 5 mega basepairs, 1 mega basepairs, 0.5 mega basepairs, 0.1 mega basepairs, 50,000 basepairs, 20,000 basepairs, or 10,000 basepairs in length.

The present invention encompasses polynucleotides with sequences complementary to those of the polynucleotides of the present invention disclosed herein. Such sequences may be complementary to the sequence disclosed as SEQ ID  
20 NO:1, the sequence contained in a deposit, and/or the nucleic acid sequence encoding the sequence disclosed as SEQ ID NO:2.

The present invention also encompasses polynucleotides capable of hybridizing, preferably under reduced stringency conditions, more preferably under stringent conditions, and most preferably under highly stringent conditions, to  
25 polynucleotides described herein. Examples of stringency conditions are shown in Table II below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

30

TABLE II

Stringency Condition	Polynucleotide Hybrid±	Hybrid Length (bp) ‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer †
A	DNA:DNA	> or equal to 50	65°C; 1xSSC – or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	< 50	Tb*; 1xSSC	Tb*; 1xSSC
C	DNA:RNA	> or equal to 50	67°C; 1xSSC – or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	< 50	Td*; 1xSSC	Td*; 1xSSC
E	RNA:RNA	> or equal to 50	70°C; 1xSSC – or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	< 50	Tf*; 1xSSC	Tf*; 1xSSC
G	DNA:DNA	> or equal to 50	65°C; 4xSSC – or- 45°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	< 50	Th*; 4xSSC	Th*; 4xSSC
I	DNA:RNA	> or equal to 50	67°C; 4xSSC – or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	< 50	Tj*; 4xSSC	Tj*; 4xSSC
K	RNA:RNA	> or equal to 50	70°C; 4xSSC – or- 40°C; 6xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	< 50	Tl*; 2xSSC	Tl*; 2xSSC
M	DNA:DNA	> or equal to 50	50°C; 4xSSC – or- 40°C 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	< 50	Tn*; 6xSSC	Tn*; 6xSSC
O	DNA:RNA	> or equal to 50	55°C; 4xSSC – or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	< 50	Tp*; 6xSSC	Tp*; 6xSSC
Q	RNA:RNA	> or equal to 50	60°C; 4xSSC – or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	< 50	Tr*; 4xSSC	Tr*; 4xSSC

‡ - The “hybrid length” is the anticipated length for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide of unknown sequence, the hybrid is assumed to be that of the hybridizing polynucleotide of the present invention. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity. Methods of aligning two or more polynucleotide sequences and/or determining the percent identity between two polynucleotide sequences are well known in the art (e.g., MegAlign program of the DNA\*Star suite of programs, etc).

† - SSPE (1xSSPE is 0.15M NaCl, 10mM NaH<sub>2</sub>PO<sub>4</sub>, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete. The hybridizations and washes may additionally include 5X Denhardt's reagent, .5-1.0% SDS, 100ug/ml denatured, fragmented salmon sperm DNA, 0.5% sodium pyrophosphate, and up to 50% formamide.

\*T<sub>b</sub> – T<sub>r</sub>: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature T<sub>m</sub> of the hybrids there T<sub>m</sub> is determined according to the following equations. For hybrids less than 18 base pairs in length,  $T_m(^{\circ}\text{C}) = 2(\# \text{ of A} + \text{T bases}) + 4(\# \text{ of G} + \text{C bases})$ . For hybrids between 18 and 49 base pairs in length,  $T_m(^{\circ}\text{C}) = 81.5 + 16.6(\log_{10}[\text{Na}^+]) + 0.41(\% \text{G} + \text{C}) - (600/\text{N})$ , where N is the number of bases in the hybrid, and [Na<sup>+</sup>] is the concentration of sodium ions in the hybridization buffer ([Na<sup>+</sup>] for 1xSSC = .165 M).

± - The present invention encompasses the substitution of any one, or more DNA or RNA hybrid partners with either a PNA, or a modified polynucleotide. Such modified polynucleotides are known in the art and are more particularly described elsewhere herein.

Additional examples of stringency conditions for polynucleotide hybridization are provided, for example, in Sambrook, J., E.F. Fritsch, and T.Maniatis, 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and Current Protocols in Molecular

Biology, 1995, F.M., Ausubel et al., eds, John Wiley and Sons, Inc., sections 2.10 and 6.3-6.4, which are hereby incorporated by reference herein.

Preferably, such hybridizing polynucleotides have at least 70% sequence identity (more preferably, at least 80% identity; and most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which they hybridize, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps. The determination of identity is well known in the art, and discussed more specifically elsewhere herein.

The invention encompasses the application of PCR methodology to the polynucleotide sequences of the present invention, the clone deposited with the ATCC, and/or the cDNA encoding the polypeptides of the present invention. PCR techniques for the amplification of nucleic acids are described in US Patent No. 4, 683, 195 and Saiki et al., Science, 239:487-491 (1988). PCR, for example, may include the following steps, of denaturation of template nucleic acid (if double-stranded), annealing of primer to target, and polymerization. The nucleic acid probed or used as a template in the amplification reaction may be genomic DNA, cDNA, RNA, or a PNA. PCR may be used to amplify specific sequences from genomic DNA, specific RNA sequence, and/or cDNA transcribed from mRNA. References for the general use of PCR techniques, including specific method parameters, include Mullis et al., Cold Spring Harbor Symp. Quant. Biol., 51:263, (1987), Ehrlich (ed), PCR Technology, Stockton Press, NY, 1989; Ehrlich et al., Science, 252:1643-1650, (1991); and "PCR Protocols, A Guide to Methods and Applications", Eds., Innis et al., Academic Press, New York, (1990).

#### Polynucleotide and Polypeptide Variants

The present invention also encompasses variants (e.g., allelic variants, orthologs, etc.) of the polynucleotide sequence disclosed herein in SEQ ID NO:1, the complementary strand thereto, and/or the cDNA sequence contained in the deposited clone.

The present invention also encompasses variants of the polypeptide sequence, and/or fragments therein, disclosed in SEQ ID NO:2, a polypeptide encoded by the

polynucleotide sequence in SEQ ID NO:1, and/or a polypeptide encoded by a cDNA in the deposited clone.

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential  
5 properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

Thus, one aspect of the invention provides an isolated nucleic acid molecule comprising, or alternatively consisting of, a polynucleotide having a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding a  
10 human phosphatase related polypeptide having an amino acid sequence as shown in the sequence listing and described in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:Z; (b) a nucleotide sequence encoding a mature human phosphatase related polypeptide having the amino acid sequence as shown in the sequence listing and described in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:Z; (c) a  
15 nucleotide sequence encoding a biologically active fragment of a human phosphatase related polypeptide having an amino acid sequence shown in the sequence listing and described in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:Z; (d) a nucleotide sequence encoding an antigenic fragment of a human phosphatase related polypeptide having an amino acid sequence shown in the sequence listing and  
20 described in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:Z; (e) a nucleotide sequence encoding a human phosphatase related polypeptide comprising the complete amino acid sequence encoded by a human cDNA plasmid contained in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:Z; (f) a nucleotide sequence encoding a mature human phosphatase related polypeptide having an amino  
25 acid sequence encoded by a human cDNA plasmid contained in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:Z; (g) a nucleotide sequence encoding a biologically active fragment of a human phosphatase related polypeptide having an amino acid sequence encoded by a human cDNA plasmid contained in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:Z; (h) a nucleotide sequence encoding an  
30 antigenic fragment of a human phosphatase related polypeptide having an amino acid sequence encoded by a human cDNA plasmid contained in SEQ ID NO:1 or the

cDNA contained in ATCC deposit No:Z; (I) a nucleotide sequence complimentary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h), above.

The present invention is also directed to polynucleotide sequences which comprise, or alternatively consist of, a polynucleotide sequence which is at least about  
5 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to, for example, any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h), above. Polynucleotides encoded by these nucleic acid molecules are also encompassed by the invention. In another embodiment, the invention encompasses nucleic acid molecules  
10 which comprise, or alternatively, consist of a polynucleotide which hybridizes under stringent conditions, or alternatively, under lower stringency conditions, to a polynucleotide in (a), (b), (c), (d), (e), (f), (g), or (h), above. Polynucleotides which hybridize to the complement of these nucleic acid molecules under stringent hybridization conditions or alternatively, under lower stringency conditions, are also  
15 encompassed by the invention, as are polypeptides encoded by these polypeptides.

Another aspect of the invention provides an isolated nucleic acid molecule comprising, or alternatively, consisting of, a polynucleotide having a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding a human phosphatase related polypeptide having an amino acid sequence as shown in  
20 the sequence listing and described in Table I; (b) a nucleotide sequence encoding a mature human phosphatase related polypeptide having the amino acid sequence as shown in the sequence listing and described in Table I; (c) a nucleotide sequence encoding a biologically active fragment of a human phosphatase related polypeptide having an amino acid sequence as shown in the sequence listing and described in Table  
25 I; (d) a nucleotide sequence encoding an antigenic fragment of a human phosphatase related polypeptide having an amino acid sequence as shown in the sequence listing and described in Table I; (e) a nucleotide sequence encoding a human phosphatase related polypeptide comprising the complete amino acid sequence encoded by a human cDNA in a cDNA plasmid contained in the ATCC Deposit and described in  
30 Table I; (f) a nucleotide sequence encoding a mature human phosphatase related polypeptide having an amino acid sequence encoded by a human cDNA in a cDNA plasmid contained in the ATCC Deposit and described in Table I; (g) a nucleotide

sequence encoding a biologically active fragment of a human phosphatase related polypeptide having an amino acid sequence encoded by a human cDNA in a cDNA plasmid contained in the ATCC Deposit and described in Table I; (h) a nucleotide sequence encoding an antigenic fragment of a human phosphatase related polypeptide having an amino acid sequence encoded by a human cDNA in a cDNA plasmid contained in the ATCC deposit and described in Table I; (i) a nucleotide sequence complimentary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h) above.

The present invention is also directed to nucleic acid molecules which comprise, or alternatively, consist of, a nucleotide sequence which is at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to, for example, any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h), above.

The present invention encompasses polypeptide sequences which comprise, or alternatively consist of, an amino acid sequence which is at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to, the following non-limited examples, the polypeptide sequence identified as SEQ ID NO:2, the polypeptide sequence encoded by a cDNA provided in the deposited clone, and/or polypeptide fragments of any of the polypeptides provided herein. Polynucleotides encoded by these nucleic acid molecules are also encompassed by the invention. In another embodiment, the invention encompasses nucleic acid molecules which comprise, or alternatively, consist of a polynucleotide which hybridizes under stringent conditions, or alternatively, under lower stringency conditions, to a polynucleotide in (a), (b), (c), (d), (e), (f), (g), or (h), above. Polynucleotides which hybridize to the complement of these nucleic acid molecules under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the invention, as are polypeptides encoded by these polypeptides.

The present invention is also directed to polypeptides which comprise, or alternatively consist of, an amino acid sequence which is at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to, for example, the

polypeptide sequence shown in SEQ ID NO:2, a polypeptide sequence encoded by the nucleotide sequence in SEQ ID NO:1, a polypeptide sequence encoded by the cDNA in cDNA plasmid:Z, and/or polypeptide fragments of any of these polypeptides (e.g., those fragments described herein). Polynucleotides which hybridize to the complement of the nucleic acid molecules encoding these polypeptides under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the present invention, as are the polypeptides encoded by these polynucleotides.

By a nucleic acid having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the nucleic acid is identical to the reference sequence except that the nucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a nucleic acid having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence referenced in Table I, the ORF (open reading frame), or any fragment specified as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least about 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the CLUSTALW computer program (Thompson, J.D., et al., Nucleic Acids Research, 2(22):4673-4680, (1994)), which is based on the algorithm of Higgins, D.G., et al., Computer Applications in the Biosciences (CABIOS), 8(2):189-191, (1992). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's.

However, the CLUSTALW algorithm automatically converts U's to T's when comparing RNA sequences to DNA sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a CLUSTALW alignment of DNA sequences to calculate percent identity via pairwise alignments are: Matrix=IUB, k-tuple=1, Number of Top Diagonals=5, Gap Penalty=3, Gap Open  
 5 Penalty 10, Gap Extension Penalty=0.1, Scoring Method=Percent, Window Size=5 or the length of the subject nucleotide sequence, whichever is shorter. For multiple alignments, the following CLUSTALW parameters are preferred: Gap Opening Penalty=10; Gap Extension Parameter=0.05; Gap Separation Penalty Range=8; End  
 10 Gap Separation Penalty=Off; % Identity for Alignment Delay=40%; Residue Specific Gaps:Off; Hydrophilic Residue Gap=Off; and Transition Weighting=0. The pairwise and multiple alignment parameters provided for CLUSTALW above represent the default parameters as provided with the AlignX software program (Vector NTI suite of programs, version 6.0).

15 The present invention encompasses the application of a manual correction to the percent identity results, in the instance where the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions. If only the local pairwise percent identity is required, no manual correction is needed. However, a manual correction may be applied to determine the global percent identity  
 20 from a global polynucleotide alignment. Percent identity calculations based upon global polynucleotide alignments are often preferred since they reflect the percent identity between the polynucleotide molecules as a whole (i.e., including any polynucleotide overhangs, not just overlapping regions), as opposed to, only local matching polynucleotides. Manual corrections for global percent identity  
 25 determinations are required since the CLUSTALW program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent  
 30 of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the CLUSTALW sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above CLUSTALW

program using the specified parameters, to arrive at a final percent identity score. This corrected score may be used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the CLUSTALW alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the CLUSTALW alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the CLUSTALW program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by CLUSTALW is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are required for the purposes of the present invention.

In addition to the above method of aligning two or more polynucleotide or polypeptide sequences to arrive at a percent identity value for the aligned sequences, it may be desirable in some circumstances to use a modified version of the CLUSTALW algorithm which takes into account known structural features of the sequences to be aligned, such as for example, the SWISS-PROT designations for each sequence. The result of such a modified CLUSTALW algorithm may provide a more accurate value of the percent identity for two polynucleotide or polypeptide sequences. Support for such a modified version of CLUSTALW is provided within the CLUSTALW algorithm and would be readily appreciated to one of skill in the art of bioinformatics.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing

alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or  
5 added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an  
10 organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level and are included in the present invention. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA  
15 technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the protein without substantial loss of biological function. The authors of Ron et al., *J. Biol. Chem.* 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8,  
20 or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein (Dobeli et al., *J. Biotechnology* 7:199-216 (1988)).

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and  
25 coworkers (*J. Biol. Chem.* 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be  
30 altered with little effect on either [binding or biological activity]." In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the protein will likely be retained when less than the majority of the residues of the protein are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Alternatively, such N-terminus or C-terminus deletions of a polypeptide of the present invention may, in fact, result in a significant increase in one or more of the biological activities of the polypeptide(s). For example, biological activity of many polypeptides are governed by the presence of regulatory domains at either one or both termini. Such regulatory domains effectively inhibit the biological activity of such polypeptides in lieu of an activation event (e.g., binding to a cognate ligand or receptor, phosphorylation, proteolytic processing, etc.). Thus, by eliminating the regulatory domain of a polypeptide, the polypeptide may effectively be rendered biologically active in the absence of an activation event.

Thus, the invention further includes polypeptide variants that show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid

substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function.

5 For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

10 As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved.

15 The invention encompasses polypeptides having a lower degree of identity but having sufficient similarity so as to perform one or more of the same functions performed by the polypeptide of the present invention. Similarity is determined by conserved amino acid substitution. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics (e.g.,  
20 chemical properties). According to Cunningham et al above, such conservative substitutions are likely to be phenotypically silent. Additional guidance concerning which amino acid changes are likely to be phenotypically silent are found in Bowie et al., Science 247:1306-1310 (1990).

25 Tolerated conservative amino acid substitutions of the present invention involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

30 In addition, the present invention also encompasses the conservative substitutions provided in Table VII below.

TABLE VII

For Amino Acid	Code	Replace with any of:
Alanine	A	D-Ala, Gly, beta-Ala, L-Cys, D-Cys
Arginine	R	D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn
Asparagine	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
Aspartic Acid	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
Cysteine	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
Glutamine	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Glutamic Acid	E	D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
Glycine	G	Ala, D-Ala, Pro, D-Pro, $\beta$ -Ala, Acp
Isoleucine	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
Leucine	L	D-Leu, Val, D-Val, Met, D-Met
Lysine	K	D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
Methionine	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val
Phenylalanine	F	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-phenylproline, cis-3,4, or 5-phenylproline
Proline	P	D-Pro, L-1-thioazolidine-4-carboxylic acid, D- or L-1-oxazolidine-4-carboxylic acid
Serine	S	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Threonine	T	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Tyrosine	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Valine	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met

Aside from the uses described above, such amino acid substitutions may also increase protein or peptide stability. The invention encompasses amino acid substitutions that contain, for example, one or more non-peptide bonds (which replace the peptide bonds) in the protein or peptide sequence. Also included are substitutions that include amino acid residues other than naturally occurring L-amino acids, e.g., D-amino acids or non-naturally occurring or synthetic amino acids, e.g.,  $\beta$  or  $\gamma$  amino acids.

Both identity and similarity can be readily calculated by reference to the following publications: Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; Informatics Computer Analysis

of Sequence Data, Part 1, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991.

5           In addition, the present invention also encompasses substitution of amino acids based upon the probability of an amino acid substitution resulting in conservation of function. Such probabilities are determined by aligning multiple genes with related function and assessing the relative penalty of each substitution to proper gene function. Such probabilities are often described in a matrix and are used  
10 by some algorithms (e.g., BLAST, CLUSTALW, GAP, etc.) in calculating percent similarity wherein similarity refers to the degree by which one amino acid may substitute for another amino acid without lose of function. An example of such a matrix is the PAM250 or BLOSUM62 matrix.

          Aside from the canonical chemically conservative substitutions referenced  
15 above, the invention also encompasses substitutions which are typically not classified as conservative, but that may be chemically conservative under certain circumstances. Analysis of enzymatic catalysis for proteases, for example, has shown that certain amino acids within the active site of some enzymes may have highly perturbed pKa's due to the unique microenvironment of the active site. Such perturbed pKa's could  
20 enable some amino acids to substitute for other amino acids while conserving enzymatic structure and function. Examples of amino acids that are known to have amino acids with perturbed pKa's are the Glu-35 residue of Lysozyme, the Ile-16 residue of Chymotrypsin, the His-159 residue of Papain, etc. The conservation of function relates to either anomalous protonation or anomalous deprotonation of such  
25 amino acids, relative to their canonical, non-perturbed pKa. The pKa perturbation may enable these amino acids to actively participate in general acid-base catalysis due to the unique ionization environment within the enzyme active site. Thus, substituting an amino acid capable of serving as either a general acid or general base within the microenvironment of an enzyme active site or cavity, as may be the case, in the same  
30 or similar capacity as the wild-type amino acid, would effectively serve as a conservative amino substitution.

Besides conservative amino acid substitution, variants of the present invention include, but are not limited to, the following: (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as, for example, an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

Moreover, the invention further includes polypeptide variants created through the application of molecular evolution ("DNA Shuffling") methodology to the polynucleotide disclosed as SEQ ID NO:1, the sequence of the clone submitted in a deposit, and/or the cDNA encoding the polypeptide disclosed as SEQ ID NO:2. Such DNA Shuffling technology is known in the art and more particularly described elsewhere herein (e.g., WPC, Stemmer, PNAS, 91:10747, (1994)), and in the Examples provided herein).

A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of the present invention having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a peptide or polypeptide

to have an amino acid sequence which comprises the amino acid sequence of the present invention, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of the present invention or  
 5 fragments thereof (e.g., the mature form and/or other fragments described herein), is 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

#### Polynucleotide and Polypeptide Fragments

The present invention is directed to polynucleotide fragments of the  
 10 polynucleotides of the invention, in addition to polypeptides encoded therein by said polynucleotides and/or fragments.

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence which: is a portion of that contained in a deposited clone, or encoding the polypeptide encoded by the cDNA in a deposited  
 15 clone; is a portion of that shown in SEQ ID NO:1 or the complementary strand thereto, or is a portion of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2. The nucleotide fragments of the invention are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt, at least about 50 nt, at least about 75  
 20 nt, or at least about 150 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in a deposited clone or the nucleotide sequence shown in SEQ ID NO:1. In this context "about" includes the particularly recited value, a value larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus, or at both termini. These  
 25 nucleotide fragments have uses that include, but are not limited to, as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments comprising, or alternatively consisting of,  
 30 a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150,

1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:1, or the complementary strand thereto, or the cDNA contained in a deposited clone. In this context "about" includes the particularly recited ranges, and ranges larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein. Also encompassed by the present invention are polynucleotides which hybridize to these nucleic acid molecules under stringent hybridization conditions or lower stringency conditions, as are the polypeptides encoded by these polynucleotides.

In the present invention, a "polypeptide fragment" refers to an amino acid sequence which is a portion of that contained in SEQ ID NO:2 or encoded by the cDNA contained in a deposited clone. Protein (polypeptide) fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments comprising, or alternatively consisting of, from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 175, 200, 225, 250, 275, or 300 amino acids in length. In this context "about" includes the particularly recited ranges or values, and ranges or values larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Preferred polypeptide fragments include the full-length protein. Further preferred polypeptide fragments include the full-length protein having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of the full-length polypeptide. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the full-length

protein. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotides encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:2 falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotides encoding these domains are also contemplated.

Other preferred polypeptide fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity. Polynucleotides encoding these polypeptide fragments are also encompassed by the invention.

In a preferred embodiment, the functional activity displayed by a polypeptide encoded by a polynucleotide fragment of the invention may be one or more biological activities typically associated with the full-length polypeptide of the invention. Illustrative of these biological activities includes the fragments ability to bind to at least one of the same antibodies which bind to the full-length protein, the fragments ability to interact with at least one of the same proteins which bind to the full-length, the fragments ability to elicit at least one of the same immune responses as the full-length protein (i.e., to cause the immune system to create antibodies specific to the same epitope, etc.), the fragments ability to bind to at least one of the same polynucleotides as the full-length protein, the fragments ability to bind to a receptor of the full-length protein, the fragments ability to bind to a ligand of the full-length protein, and the fragments ability to multimerize with the full-length protein. However, the skilled artisan would appreciate that some fragments may have biological activities which are desirable and directly inapposite to the biological

activity of the full-length protein. The functional activity of polypeptides of the invention, including fragments, variants, derivatives, and analogs thereof can be determined by numerous methods available to the skilled artisan, some of which are described elsewhere herein.

5           The present invention encompasses polypeptides comprising, or alternatively consisting of, an epitope of the polypeptide having an amino acid sequence of SEQ ID NO:2, or an epitope of the polypeptide sequence encoded by a polynucleotide sequence contained in ATCC deposit No. Z or encoded by a polynucleotide that hybridizes to the complement of the sequence of SEQ ID NO:1 or contained in ATCC  
10       deposit No. Z under stringent hybridization conditions or lower stringency hybridization conditions as defined supra. The present invention further encompasses polynucleotide sequences encoding an epitope of a polypeptide sequence of the invention (such as, for example, the sequence disclosed in SEQ ID NO:1), polynucleotide sequences of the complementary strand of a polynucleotide sequence  
15       encoding an epitope of the invention, and polynucleotide sequences which hybridize to the complementary strand under stringent hybridization conditions or lower stringency hybridization conditions defined supra.

          The term "epitopes," as used herein, refers to portions of a polypeptide having antigenic or immunogenic activity in an animal, preferably a mammal, and most  
20       preferably in a human. In a preferred embodiment, the present invention encompasses a polypeptide comprising an epitope, as well as the polynucleotide encoding this polypeptide. An "immunogenic epitope," as used herein, is defined as a portion of a protein that elicits an antibody response in an animal, as determined by any method known in the art, for example, by the methods for generating antibodies described  
25       infra. (See, for example, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983)). The term "antigenic epitope," as used herein, is defined as a portion of a protein to which an antibody can immunospecifically bind its antigen as determined by any method well known in the art, for example, by the immunoassays described herein. Immunospecific binding excludes non-specific binding but does not  
30       necessarily exclude cross- reactivity with other antigens. Antigenic epitopes need not necessarily be immunogenic.

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985), further described in U.S. Patent No. 4,631,211).

In the present invention, antigenic epitopes preferably contain a sequence of at  
5 least 4, at least 5, at least 6, at least 7, more preferably at least 8, at least 9, at least 10,  
at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at  
least 30, at least 40, at least 50, and, most preferably, between about 15 to about 30  
amino acids. Preferred polypeptides comprising immunogenic or antigenic epitopes  
are at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100  
10 amino acid residues in length, or longer. Additional non-exclusive preferred antigenic  
epitopes include the antigenic epitopes disclosed herein, as well as portions thereof.  
Antigenic epitopes are useful, for example, to raise antibodies, including monoclonal  
antibodies, that specifically bind the epitope. Preferred antigenic epitopes include the  
antigenic epitopes disclosed herein, as well as any combination of two, three, four,  
15 five or more of these antigenic epitopes. Antigenic epitopes can be used as the target  
molecules in immunoassays. (See, for instance, Wilson et al., Cell 37:767-778 (1984);  
Sutcliffe et al., Science 219:660-666 (1983)).

Similarly, immunogenic epitopes can be used, for example, to induce  
antibodies according to methods well known in the art. (See, for instance, Sutcliffe et  
20 al., supra; Wilson et al., supra; Chow et al., Proc. Natl. Acad. Sci. USA 82:910-914;  
and Bittle et al., J. Gen. Virol. 66:2347-2354 (1985). Preferred immunogenic epitopes  
include the immunogenic epitopes disclosed herein, as well as any combination of  
two, three, four, five or more of these immunogenic epitopes. The polypeptides  
comprising one or more immunogenic epitopes may be presented for eliciting an  
25 antibody response together with a carrier protein, such as an albumin, to an animal  
system (such as rabbit or mouse), or, if the polypeptide is of sufficient length (at least  
about 25 amino acids), the polypeptide may be presented without a carrier. However,  
immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to  
be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes  
30 in a denatured polypeptide (e.g., in Western blotting).

Epitope-bearing polypeptides of the present invention may be used to induce  
antibodies according to methods well known in the art including, but not limited to, in

vivo immunization, in vitro immunization, and phage display methods. See, e.g., Sutcliffe et al., supra; Wilson et al., supra, and Bittle et al., J. Gen. Virol., 66:2347-2354 (1985). If in vivo immunization is used, animals may be immunized with free peptide; however, anti-peptide antibody titer may be boosted by coupling the peptide to a macromolecular carrier, such as keyhole limpet hemacyanin (KLH) or tetanus toxoid. For instance, peptides containing cysteine residues may be coupled to a carrier using a linker such as maleimidobenzoyl- N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carriers using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice are immunized with either free or carrier- coupled peptides, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 µg of peptide or carrier protein and Freund's adjuvant or any other adjuvant known for stimulating an immune response. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of anti-peptide antibody which can be detected, for example, by ELISA assay using free peptide adsorbed to a solid surface. The titer of anti-peptide antibodies in serum from an immunized animal may be increased by selection of anti-peptide antibodies, for instance, by adsorption to the peptide on a solid support and elution of the selected antibodies according to methods well known in the art.

As one of skill in the art will appreciate, and as discussed above, the polypeptides of the present invention comprising an immunogenic or antigenic epitope can be fused to other polypeptide sequences. For example, the polypeptides of the present invention may be fused with the constant domain of immunoglobulins (IgA, IgE, IgG, IgM), or portions thereof (CH1, CH2, CH3, or any combination thereof and portions thereof) resulting in chimeric polypeptides. Such fusion proteins may facilitate purification and may increase half-life in vivo. This has been shown for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. See, e.g., EP 394,827; Traunecker et al., Nature, 331:84-86 (1988). Enhanced delivery of an antigen across the epithelial barrier to the immune system has been demonstrated for antigens (e.g., insulin) conjugated to an FcRn binding partner such as IgG or Fc fragments (see, e.g., PCT Publications WO

96/22024 and WO 99/04813). IgG Fusion proteins that have a disulfide-linked dimeric structure due to the IgG portion disulfide bonds have also been found to be more efficient in binding and neutralizing other molecules than monomeric polypeptides or fragments thereof alone. See, e.g., Fountoulakis et al., *J. Biochem.*, 5 270:3958-3964 (1995). Nucleic acids encoding the above epitopes can also be recombined with a gene of interest as an epitope tag (e.g., the hemagglutinin ("HA") tag or flag tag) to aid in detection and purification of the expressed polypeptide. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht et al., 1991, 10 *Proc. Natl. Acad. Sci. USA* 88:8972- 897). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the open reading frame of the gene is translationally fused to an amino-terminal tag consisting of six histidine residues. The tag serves as a matrix binding domain for the fusion protein. Extracts from cells infected with the recombinant vaccinia virus are loaded onto Ni<sup>2+</sup> 15 nitriloacetic acid-agarose column and histidine-tagged proteins can be selectively eluted with imidazole-containing buffers.

Additional fusion proteins of the invention may be generated through the techniques of gene-shuffling, motif-shuffling, exon-shuffling, and/or codon-shuffling (collectively referred to as "DNA shuffling"). DNA shuffling may be employed to 20 modulate the activities of polypeptides of the invention, such methods can be used to generate polypeptides with altered activity, as well as agonists and antagonists of the polypeptides. See, generally, U.S. Patent Nos. 5,605,793; 5,811,238; 5,830,721; 5,834,252; and 5,837,458, and Patten et al., *Curr. Opinion Biotechnol.* 8:724-33 (1997); Harayama, *Trends Biotechnol.* 16(2):76-82 (1998); Hansson, et al., *J. Mol.* 25 *Biol.* 287:265-76 (1999); and Lorenzo and Blasco, *Biotechniques* 24(2):308- 13 (1998) (each of these patents and publications are hereby incorporated by reference in its entirety). In one embodiment, alteration of polynucleotides corresponding to SEQ ID NO:1 and the polypeptides encoded by these polynucleotides may be achieved by DNA shuffling. DNA shuffling involves the assembly of two or more DNA segments 30 by homologous or site-specific recombination to generate variation in the polynucleotide sequence. In another embodiment, polynucleotides of the invention, or the encoded polypeptides, may be altered by being subjected to random mutagenesis

by error-prone PCR, random nucleotide insertion or other methods prior to recombination. In another embodiment, one or more components, motifs, sections, parts, domains, fragments, etc., of a polynucleotide encoding a polypeptide of the invention may be recombined with one or more components, motifs, sections, parts,  
 5 domains, fragments, etc. of one or more heterologous molecules.

#### Antibodies

Further polypeptides of the invention relate to antibodies and T-cell antigen receptors (TCR) which immunospecifically bind a polypeptide, polypeptide fragment, or variant of SEQ ID NO:Y, and/or an epitope, of the present invention (as  
 10 determined by immunoassays well known in the art for assaying specific antibody-antigen binding). Antibodies of the invention include, but are not limited to, polyclonal, monoclonal, monovalent, bispecific, heteroconjugate, multispecific, human, humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab') fragments, fragments produced by a Fab expression library, anti-idiotypic  
 15 (anti-Id) antibodies (including, e.g., anti-Id antibodies to antibodies of the invention), and epitope-binding fragments of any of the above. The term "antibody," as used herein, refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that immunospecifically binds an antigen. The immunoglobulin molecules of the invention  
 20 can be of any type (e.g., IgG, IgE, IgM, IgD, IgA and IgY), class (e.g., IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2) or subclass of immunoglobulin molecule. Moreover, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules, as well as, antibody fragments (such as, for example, Fab and F(ab')<sub>2</sub> fragments) which are capable of specifically binding to protein. Fab and F(ab')<sub>2</sub>  
 25 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation of the animal or plant, and may have less non-specific tissue binding than an intact antibody (Wahl et al., J. Nucl. Med... 24:316-325 (1983)). Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric,  
 30 single chain, and humanized antibodies.

Most preferably the antibodies are human antigen-binding antibody fragments of the present invention and include, but are not limited to, Fab, Fab' and F(ab')<sub>2</sub>, Fd,

single-chain Fvs (scFv), single-chain antibodies, disulfide-linked Fvs (sdFv) and fragments comprising either a VL or VH domain. Antigen-binding antibody fragments, including single-chain antibodies, may comprise the variable region(s) alone or in combination with the entirety or a portion of the following: hinge region,  
5 CH1, CH2, and CH3 domains. Also included in the invention are antigen-binding fragments also comprising any combination of variable region(s) with a hinge region, CH1, CH2, and CH3 domains. The antibodies of the invention may be from any animal origin including birds and mammals. Preferably, the antibodies are human, murine (e.g., mouse and rat), donkey, ship rabbit, goat, guinea pig, camel, horse, or  
10 chicken. As used herein, "human" antibodies include antibodies having the amino acid sequence of a human immunoglobulin and include antibodies isolated from human immunoglobulin libraries or from animals transgenic for one or more human immunoglobulin and that do not express endogenous immunoglobulins, as described infra and, for example in, U.S. Patent No. 5,939,598 by Kucherlapati et al.

15 The antibodies of the present invention may be monospecific, bispecific, trispecific or of greater multispecificity. Multispecific antibodies may be specific for different epitopes of a polypeptide of the present invention or may be specific for both a polypeptide of the present invention as well as for a heterologous epitope, such as a heterologous polypeptide or solid support material. See, e.g., PCT publications WO  
20 93/17715; WO 92/08802; WO 91/00360; WO 92/05793; Tutt, et al., J. Immunol. 147:60-69 (1991); U.S. Patent Nos. 4,474,893; 4,714,681; 4,925,648; 5,573,920; 5,601,819; Kostelny et al., J. Immunol. 148:1547-1553 (1992).

Antibodies of the present invention may be described or specified in terms of the epitope(s) or portion(s) of a polypeptide of the present invention which they  
25 recognize or specifically bind. The epitope(s) or polypeptide portion(s) may be specified as described herein, e.g., by N-terminal and C-terminal positions, by size in contiguous amino acid residues, or listed in the Tables and Figures. Antibodies which specifically bind any epitope or polypeptide of the present invention may also be excluded. Therefore, the present invention includes antibodies that specifically bind  
30 polypeptides of the present invention, and allows for the exclusion of the same.

Antibodies of the present invention may also be described or specified in terms of their cross-reactivity. Antibodies that do not bind any other analog, ortholog,

or homologue of a polypeptide of the present invention are included. Antibodies that bind polypeptides with at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55%, and at least 50% identity (as calculated using methods known in the art and described herein) to a polypeptide of the present invention are also included in the present invention. In specific embodiments, antibodies of the present invention cross-react with murine, rat and/or rabbit homologues of human proteins and the corresponding epitopes thereof. Antibodies that do not bind polypeptides with less than 95%, less than 90%, less than 85%, less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, and less than 50% identity (as calculated using methods known in the art and described herein) to a polypeptide of the present invention are also included in the present invention. In a specific embodiment, the above-described cross-reactivity is with respect to any single specific antigenic or immunogenic polypeptide, or combination(s) of 2, 3, 4, 5, or more of the specific antigenic and/or immunogenic polypeptides disclosed herein. Further included in the present invention are antibodies which bind polypeptides encoded by polynucleotides which hybridize to a polynucleotide of the present invention under stringent hybridization conditions (as described herein). Antibodies of the present invention may also be described or specified in terms of their binding affinity to a polypeptide of the invention. Preferred binding affinities include those with a dissociation constant or  $K_d$  less than  $5 \times 10^{-2}$  M,  $10^{-2}$  M,  $5 \times 10^{-3}$  M,  $10^{-3}$  M,  $5 \times 10^{-4}$  M,  $10^{-4}$  M,  $5 \times 10^{-5}$  M,  $10^{-5}$  M,  $5 \times 10^{-6}$  M,  $10^{-6}$  M,  $5 \times 10^{-7}$  M,  $10^{-7}$  M,  $5 \times 10^{-8}$  M,  $10^{-8}$  M,  $5 \times 10^{-9}$  M,  $10^{-9}$  M,  $5 \times 10^{-10}$  M,  $10^{-10}$  M,  $5 \times 10^{-11}$  M,  $10^{-11}$  M,  $5 \times 10^{-12}$  M,  $10^{-12}$  M,  $5 \times 10^{-13}$  M,  $10^{-13}$  M,  $5 \times 10^{-14}$  M,  $10^{-14}$  M,  $5 \times 10^{-15}$  M, or  $10^{-15}$  M.

The invention also provides antibodies that competitively inhibit binding of an antibody to an epitope of the invention as determined by any method known in the art for determining competitive binding, for example, the immunoassays described herein. In preferred embodiments, the antibody competitively inhibits binding to the epitope by at least 95%, at least 90%, at least 85 %, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50%.

Antibodies of the present invention may act as agonists or antagonists of the polypeptides of the present invention. For example, the present invention includes

antibodies which disrupt the receptor/ligand interactions with the polypeptides of the invention either partially or fully. Preferably, antibodies of the present invention bind an antigenic epitope disclosed herein, or a portion thereof. The invention features both receptor-specific antibodies and ligand-specific antibodies. The invention also  
5 features receptor-specific antibodies which do not prevent ligand binding but prevent receptor activation. Receptor activation (i.e., signaling) may be determined by techniques described herein or otherwise known in the art. For example, receptor activation can be determined by detecting the phosphorylation (e.g., tyrosine or serine/threonine) of the receptor or its substrate by immunoprecipitation followed by  
10 western blot analysis (for example, as described supra). In specific embodiments, antibodies are provided that inhibit ligand activity or receptor activity by at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50% of the activity in absence of the antibody.

The invention also features receptor-specific antibodies which both prevent  
15 ligand binding and receptor activation as well as antibodies that recognize the receptor-ligand complex, and, preferably, do not specifically recognize the unbound receptor or the unbound ligand. Likewise, included in the invention are neutralizing antibodies which bind the ligand and prevent binding of the ligand to the receptor, as well as antibodies which bind the ligand, thereby preventing receptor activation, but  
20 do not prevent the ligand from binding the receptor. Further included in the invention are antibodies which activate the receptor. These antibodies may act as receptor agonists, i.e., potentiate or activate either all or a subset of the biological activities of the ligand-mediated receptor activation, for example, by inducing dimerization of the receptor. The antibodies may be specified as agonists, antagonists or inverse agonists  
25 for biological activities comprising the specific biological activities of the peptides of the invention disclosed herein. The above antibody agonists can be made using methods known in the art. See, e.g., PCT publication WO 96/40281; U.S. Patent No. 5,811,097; Deng et al., Blood 92(6):1981-1988 (1998); Chen et al., Cancer Res. 58(16):3668-3678 (1998); Harrop et al., J. Immunol. 161(4):1786-1794 (1998); Zhu et al., Cancer Res. 58(15):3209-3214 (1998); Yoon et al., J. Immunol. 160(7):3170-3179  
30 (1998); Prat et al., J. Cell. Sci. 111(Pt2):237-247 (1998); Pitard et al., J. Immunol. Methods 205(2):177-190 (1997); Liautard et al., Cytokine 9(4):233-241 (1997);

Carlson et al., J. Biol. Chem. 272(17):11295-11301 (1997); Taryman et al., Neuron 14(4):755-762 (1995); Muller et al., Structure 6(9):1153-1167 (1998); Bartunek et al., Cytokine 8(1):14-20 (1996) (which are all incorporated by reference herein in their entireties).

5           Antibodies of the present invention may be used, for example, but not limited to, to purify, detect, and target the polypeptides of the present invention, including both in vitro and in vivo diagnostic and therapeutic methods. For example, the antibodies have use in immunoassays for qualitatively and quantitatively measuring levels of the polypeptides of the present invention in biological samples. See, e.g.,  
10   Harlow et al., Antibodies: A Laboratory Manual, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988) (incorporated by reference herein in its entirety).

          As discussed in more detail below, the antibodies of the present invention may be used either alone or in combination with other compositions. The antibodies may further be recombinantly fused to a heterologous polypeptide at the N- or C-terminus  
15   or chemically conjugated (including covalently and non-covalently conjugations) to polypeptides or other compositions. For example, antibodies of the present invention may be recombinantly fused or conjugated to molecules useful as labels in detection assays and effector molecules such as heterologous polypeptides, drugs, radionucleotides, or toxins. See, e.g., PCT publications WO 92/08495; WO 91/14438;  
20   WO 89/12624; U.S. Patent No. 5,314,995; and EP 396,387.

          The antibodies of the invention include derivatives that are modified, i.e., by the covalent attachment of any type of molecule to the antibody such that covalent attachment does not prevent the antibody from generating an anti-idiotypic response. For example, but not by way of limitation, the antibody derivatives include antibodies  
25   that have been modified, e.g., by glycosylation, acetylation, pegylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by known techniques, including, but not limited to specific chemical cleavage, acetylation, formylation, metabolic  
30   synthesis of tunicamycin, etc. Additionally, the derivative may contain one or more non-classical amino acids.

The antibodies of the present invention may be generated by any suitable method known in the art.

The antibodies of the present invention may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan (Harlow, et al., *Antibodies: A Laboratory Manual*, (Cold spring Harbor Laboratory Press, 2<sup>nd</sup> ed. (1988); and *Current Protocols*, Chapter 2; which are hereby incorporated herein by reference in its entirety). In a preferred method, a preparation of the BMY\_HPP13 protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity. For example, a polypeptide of the invention can be administered to various host animals including, but not limited to, rabbits, mice, rats, etc. to induce the production of sera containing polyclonal antibodies specific for the antigen. The administration of the polypeptides of the present invention may entail one or more injections of an immunizing agent and, if desired, an adjuvant. Various adjuvants may be used to increase the immunological response, depending on the host species, and include but are not limited to, Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and corynebacterium parvum. Such adjuvants are also well known in the art. For the purposes of the invention, "immunizing agent" may be defined as a polypeptide of the invention, including fragments, variants, and/or derivatives thereof, in addition to fusions with heterologous polypeptides and other forms of the polypeptides described herein.

Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections, though they may also be given intramuscularly, and/or through IV). The immunizing agent may include polypeptides of the present invention or a fusion protein or variants thereof. Depending upon the nature of the polypeptides (i.e., percent hydrophobicity, percent hydrophilicity, stability, net charge, isoelectric point etc.), it may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Such conjugation includes either chemical conjugation by

derivitizing active chemical functional groups to both the polypeptide of the present invention and the immunogenic protein such that a covalent bond is formed, or through fusion-protein based methodology, or other methods known to the skilled artisan. Examples of such immunogenic proteins include, but are not limited to  
5 keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions,  
10 keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum. Additional examples of adjuvants which may be employed includes the MPL-TDM adjuvant (monophosphoryl lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

15 The antibodies of the present invention may comprise monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975) and U.S. Pat. No. 4,376,110, by Harlow, et al., Antibodies: A Laboratory Manual, (Cold spring Harbor Laboratory Press, 2<sup>nd</sup> ed. (1988), by Hammerling, et al., Monoclonal Antibodies and  
20 T-Cell Hybridomas (Elsevier, N.Y., pp. 563-681 (1981); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976), or other methods known to the artisan. Other examples of methods which may be employed for producing monoclonal antibodies includes, but are not limited to, the human B-cell hybridoma technique (Kosbor et al., 1983, Immunology Today 4:72; Cole et al.,  
25 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030), and the EBV-hybridoma technique (Cole et al., 1985, Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo. Production of high titers of mAbs in  
30 vivo makes this the presently preferred method of production.

In a hybridoma method, a mouse, a humanized mouse, a mouse with a human immune system, hamster, or other appropriate host animal, is typically immunized

with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized in vitro.

The immunizing agent will typically include polypeptides of the present invention or a fusion protein thereof. Preferably, the immunizing agent consists of an  
5     BMV\_HPP13 polypeptide or, more preferably, with a BMV\_HPP13 polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56 degrees C), and  
10    supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 ug/ml of streptomycin. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing  
15    agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986), pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture  
20    medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-  
25    deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute  
30    Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. More preferred are the parent myeloma cell line (SP2O) as provided by the ATCC. As inferred throughout the specification, human

myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

5       The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the polypeptides of the present invention. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked  
10 immunoabsorbant assay (ELISA). Such techniques are known in the art and within the skill of the artisan. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollart, Anal. Biochem., 107:220 (1980).

      After the desired hybridoma cells are identified, the clones may be subcloned  
15 by limiting dilution procedures and grown by standard methods (Goding, supra, and/or according to Wands et al. (Gastroenterology 80:225-232 (1981)). Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640. Alternatively, the hybridoma cells may be grown in vivo as ascites in a mammal.

20       The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-sepharose, hydroxyapatite chromatography, gel exclusion chromatography, gel electrophoresis, dialysis, or affinity chromatography.

25       The skilled artisan would acknowledge that a variety of methods exist in the art for the production of monoclonal antibodies and thus, the invention is not limited to their sole production in hybridomas. For example, the monoclonal antibodies may be made by recombinant DNA methods, such as those described in US patent No. 4, 816, 567. In this context, the term "monoclonal antibody" refers to an antibody  
30 derived from a single eukaryotic, phage, or prokaryotic clone. The DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of

binding specifically to genes encoding the heavy and light chains of murine antibodies, or such chains from human, humanized, or other sources). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transformed into host cells such as Simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (US Patent No. 4, 816, 567; Morrison et al, supra) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art. Monoclonal antibodies can be prepared using a wide variety of techniques known in the art including the use of hybridoma, recombinant, and phage display technologies, or a combination thereof. For example, monoclonal antibodies can be produced using hybridoma techniques including those known in the art and taught, for example, in Harlow et al., *Antibodies: A Laboratory Manual*, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, et al., in: *Monoclonal Antibodies and T-Cell Hybridomas* 563-681 (Elsevier, N.Y., 1981) (said references incorporated by reference in their entireties). The term "monoclonal antibody" as used herein is not limited to antibodies produced

through hybridoma technology. The term "monoclonal antibody" refers to an antibody that is derived from a single clone, including any eukaryotic, prokaryotic, or phage clone, and not the method by which it is produced.

5 Methods for producing and screening for specific antibodies using hybridoma technology are routine and well known in the art and are discussed in detail in the Examples described herein. In a non-limiting example, mice can be immunized with a polypeptide of the invention or a cell expressing such peptide. Once an immune response is detected, e.g., antibodies specific for the antigen are detected in the mouse serum, the mouse spleen is harvested and splenocytes isolated. The splenocytes are  
10 then fused by well known techniques to any suitable myeloma cells, for example cells from cell line SP20 available from the ATCC. Hybridomas are selected and cloned by limited dilution. The hybridoma clones are then assayed by methods known in the art for cells that secrete antibodies capable of binding a polypeptide of the invention. Ascites fluid, which generally contains high levels of antibodies, can be generated by  
15 immunizing mice with positive hybridoma clones.

Accordingly, the present invention provides methods of generating monoclonal antibodies as well as antibodies produced by the method comprising culturing a hybridoma cell secreting an antibody of the invention wherein, preferably, the hybridoma is generated by fusing splenocytes isolated from a mouse immunized  
20 with an antigen of the invention with myeloma cells and then screening the hybridomas resulting from the fusion for hybridoma clones that secrete an antibody able to bind a polypeptide of the invention.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, Fab and F(ab')<sub>2</sub> fragments of the invention may be  
25 produced by proteolytic cleavage of immunoglobulin molecules, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')<sub>2</sub> fragments). F(ab')<sub>2</sub> fragments contain the variable region, the light chain constant region and the CH1 domain of the heavy chain.

For example, the antibodies of the present invention can also be generated  
30 using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of phage particles which carry the polynucleotide sequences encoding them. In a particular embodiment, such

phage can be utilized to display antigen binding domains expressed from a repertoire or combinatorial antibody library (e.g., human or murine). Phage expressing an antigen binding domain that binds the antigen of interest can be selected or identified with antigen, e.g., using labeled antigen or antigen bound or captured to a solid surface or bead. Phage used in these methods are typically filamentous phage including fd and M13 binding domains expressed from phage with Fab, Fv or disulfide stabilized Fv antibody domains recombinantly fused to either the phage gene III or gene VIII protein. Examples of phage display methods that can be used to make the antibodies of the present invention include those disclosed in Brinkman et al., J. Immunol. Methods 182:41-50 (1995); Ames et al., J. Immunol. Methods 184:177-186 (1995); Kettleborough et al., Eur. J. Immunol. 24:952-958 (1994); Persic et al., Gene 187 9-18 (1997); Burton et al., Advances in Immunology 57:191-280 (1994); PCT application No. PCT/GB91/01134; PCT publications WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and U.S. Patent Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,908; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108; each of which is incorporated herein by reference in its entirety.

As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, e.g., as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and F(ab')<sub>2</sub> fragments can also be employed using methods known in the art such as those disclosed in PCT publication WO 92/22324; Mullinax et al., BioTechniques 12(6):864-869 (1992); and Sawai et al., AJRI 34:26-34 (1995); and Better et al., Science 240:1041-1043 (1988) (said references incorporated by reference in their entireties). Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. Patents 4,946,778 and 5,258,498; Huston et al., Methods in Enzymology 203:46-88 (1991); Shu et al., PNAS 90:7995-7999 (1993); and Skerra et al., Science 240:1038-1040 (1988).

For some uses, including in vivo use of antibodies in humans and in vitro detection assays, it may be preferable to use chimeric, humanized, or human antibodies. A chimeric antibody is a molecule in which different portions of the antibody are derived from different animal species, such as antibodies having a  
5 variable region derived from a murine monoclonal antibody and a human immunoglobulin constant region. Methods for producing chimeric antibodies are known in the art. See e.g., Morrison, *Science* 229:1202 (1985); Oi et al., *BioTechniques* 4:214 (1986); Gillies et al., (1989) *J. Immunol. Methods* 125:191-202; Cabilly et al., Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., *Nature* 312:643  
10 (1984); Neuberger et al., *Nature* 314:268 (1985); U.S. Patent Nos. 5,807,715; 4,816,567; and 4,816,397, which are incorporated herein by reference in their entirety. Humanized antibodies are antibody molecules from non-human species antibody that binds the desired antigen having one or more complementarity determining regions  
15 (CDRs) from the non-human species and a framework regions from a human immunoglobulin molecule. Often, framework residues in the human framework regions will be substituted with the corresponding residue from the CDR donor antibody to alter, preferably improve, antigen binding. These framework substitutions are identified by methods well known in the art, e.g., by modeling of the interactions  
20 of the CDR and framework residues to identify framework residues important for antigen binding and sequence comparison to identify unusual framework residues at particular positions. (See, e.g., Queen et al., U.S. Patent No. 5,585,089; Riechmann et al., *Nature* 332:323 (1988), which are incorporated herein by reference in their entirety.) Antibodies can be humanized using a variety of techniques known in the  
25 art including, for example, CDR-grafting (EP 239,400; PCT publication WO 91/09967; U.S. Patent Nos. 5,225,539; 5,530,101; and 5,585,089), veneering or resurfacing (EP 592,106; EP 519,596; Padlan, *Molecular Immunology* 28(4/5):489-498 (1991); Studnicka et al., *Protein Engineering* 7(6):805-814 (1994); Roguska. et al., *PNAS* 91:969-973 (1994)), and chain shuffling (U.S. Patent No. 5,565,332).  
30 Generally, a humanized antibody has one or more amino acid residues introduced into it from a source that is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable

domain. Humanization can be essentially performed following the methods of Winter and co-workers (Jones et al., *Nature*, 321:522-525 (1986); Reichmann et al., *Nature*, 332:323-327 (1988); Verhoeyen et al., *Science*, 239:1534-1536 (1988), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (US Patent No. 4, 816, 567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possible some FR residues are substituted from analogous sites in rodent antibodies.

In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., *Nature*, 321:522-525 (1986); Reichmann et al., *Nature* 332:323-329 (1988) and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992).

Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Human antibodies can be made by a variety of methods known in the art including phage display methods described above using antibody libraries derived from human immunoglobulin sequences. See also, U.S. Patent Nos. 4,444,887 and 4,716,111; and PCT publications WO 98/46645, WO 98/50433, WO 98/24893, WO 98/16654, WO 96/34096, WO 96/33735, and WO 91/10741; each of which is incorporated herein by reference in its entirety. The techniques of Cole et al., and Boerder et al., are also available for the preparation of human monoclonal antibodies (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Riss, (1985); and Boerner et al., *J. Immunol.*, 147(1):86-95, (1991)).

Human antibodies can also be produced using transgenic mice which are incapable of expressing functional endogenous immunoglobulins, but which can express human immunoglobulin genes. For example, the human heavy and light chain immunoglobulin gene complexes may be introduced randomly or by homologous

recombination into mouse embryonic stem cells. Alternatively, the human variable region, constant region, and diversity region may be introduced into mouse embryonic stem cells in addition to the human heavy and light chain genes. The mouse heavy and light chain immunoglobulin genes may be rendered non-functional separately or  
 5 simultaneously with the introduction of human immunoglobulin loci by homologous recombination. In particular, homozygous deletion of the JH region prevents endogenous antibody production. The modified embryonic stem cells are expanded and microinjected into blastocysts to produce chimeric mice. The chimeric mice are then bred to produce homozygous offspring which express human antibodies. The  
 10 transgenic mice are immunized in the normal fashion with a selected antigen, e.g., all or a portion of a polypeptide of the invention. Monoclonal antibodies directed against the antigen can be obtained from the immunized, transgenic mice using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo  
 15 class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA, IgM and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar, *Int. Rev. Immunol.* 13:65-93 (1995). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such  
 20 antibodies, see, e.g., PCT publications WO 98/24893; WO 92/01047; WO 96/34096; WO 96/33735; European Patent No. 0 598 877; U.S. Patent Nos. 5,413,923; 5,625,126; 5,633,425; 5,569,825; 5,661,016; 5,545,806; 5,814,318; 5,885,793; 5,916,771; and 5,939,598, which are incorporated by reference herein in their entirety. In addition, companies such as Abgenix, Inc. (Freemont, CA), Genpharm (San Jose,  
 25 CA), and Medarex, Inc. (Princeton, NJ) can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge,  
 30 human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and creation of an antibody repertoire. This approach is described, for example, in US patent Nos. 5,545,807;

5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,106, and in the following scientific publications: Marks et al., *Biotechnol.*, 10:779-783 (1992); Lonberg et al., *Nature* 368:856-859 (1994); Fishwild et al., *Nature Biotechnol.*, 14:845-51 (1996); Neuberger, *Nature Biotechnol.*, 14:826 (1996); Lonberg and Huszer, *Intern. Rev. Immunol.*, 13:65-93 (1995).

Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, e.g., a mouse antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers et al., *Bio/technology* 12:899-903 (1988)).

Further, antibodies to the polypeptides of the invention can, in turn, be utilized to generate anti-idiotypic antibodies that "mimic" polypeptides of the invention using techniques well known to those skilled in the art. (See, e.g., Greenspan & Bona, *FASEB J.* 7(5):437-444; (1989) and Nissinoff, *J. Immunol.* 147(8):2429-2438 (1991)). For example, antibodies which bind to and competitively inhibit polypeptide multimerization and/or binding of a polypeptide of the invention to a ligand can be used to generate anti-idiotypes that "mimic" the polypeptide multimerization and/or binding domain and, as a consequence, bind to and neutralize polypeptide and/or its ligand. Such neutralizing anti-idiotypes or Fab fragments of such anti-idiotypes can be used in therapeutic regimens to neutralize polypeptide ligand. For example, such anti-idiotypic antibodies can be used to bind a polypeptide of the invention and/or to bind its ligands/receptors, and thereby block its biological activity.

Such anti-idiotypic antibodies capable of binding to the BMY\_HPP13 polypeptide can be produced in a two-step procedure. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody that binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones that produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can

be used to immunize an animal to induce formation of further protein-specific antibodies.

The antibodies of the present invention may be bispecific antibodies. Bispecific antibodies are monoclonal, Preferably human or humanized, antibodies that  
5 have binding specificities for at least two different antigens. In the present invention, one of the binding specificities may be directed towards a polypeptide of the present invention, the other may be for any other antigen, and preferably for a cell-surface protein, receptor, receptor subunit, tissue-specific antigen, virally derived protein, virally encoded envelope protein, bacterially derived protein, or bacterial surface  
10 protein, etc.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, *Nature*, 305:537-539 (1983). Because of  
15 the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker  
20 et al., *EMBO J.*, 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have  
25 the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transformed into a suitable host organism. For further details of generating bispecific antibodies see, for example  
30 Suresh et al., *Meth. In Enzym.*, 121:210 (1986).

Heteroconjugate antibodies are also contemplated by the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such

antibodies have, for example, been proposed to target immune system cells to unwanted cells (US Patent No. 4, 676, 980), and for the treatment of HIV infection (WO 91/00360; WO 92/20373; and EP03089). It is contemplated that the antibodies may be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioester bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in US Patent No. 4,676,980.

#### Polynucleotides Encoding Antibodies

The invention further provides polynucleotides comprising a nucleotide sequence encoding an antibody of the invention and fragments thereof. The invention also encompasses polynucleotides that hybridize under stringent or lower stringency hybridization conditions, e.g., as defined supra, to polynucleotides that encode an antibody, preferably, that specifically binds to a polypeptide of the invention, preferably, an antibody that binds to a polypeptide having the amino acid sequence of SEQ ID NO:Y.

The polynucleotides may be obtained, and the nucleotide sequence of the polynucleotides determined, by any method known in the art. For example, if the nucleotide sequence of the antibody is known, a polynucleotide encoding the antibody may be assembled from chemically synthesized oligonucleotides (e.g., as described in Kutmeier et al., *BioTechniques* 17:242 (1994)), which, briefly, involves the synthesis of overlapping oligonucleotides containing portions of the sequence encoding the antibody, annealing and ligating of those oligonucleotides, and then amplification of the ligated oligonucleotides by PCR.

Alternatively, a polynucleotide encoding an antibody may be generated from nucleic acid from a suitable source. If a clone containing a nucleic acid encoding a particular antibody is not available, but the sequence of the antibody molecule is known, a nucleic acid encoding the immunoglobulin may be chemically synthesized or obtained from a suitable source (e.g., an antibody cDNA library, or a cDNA library generated from, or nucleic acid, preferably poly A+ RNA, isolated from, any tissue or cells expressing the antibody, such as hybridoma cells selected to express an antibody of the invention) by PCR amplification using synthetic primers hybridizable to the 3'

and 5' ends of the sequence or by cloning using an oligonucleotide probe specific for the particular gene sequence to identify, e.g., a cDNA clone from a cDNA library that encodes the antibody. Amplified nucleic acids generated by PCR may then be cloned into replicable cloning vectors using any method well known in the art.

5           Once the nucleotide sequence and corresponding amino acid sequence of the antibody is determined, the nucleotide sequence of the antibody may be manipulated using methods well known in the art for the manipulation of nucleotide sequences, e.g., recombinant DNA techniques, site directed mutagenesis, PCR, etc. (see, for example, the techniques described in Sambrook et al., 1990, *Molecular Cloning, A Laboratory Manual*, 2d Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY and Ausubel et al., eds., 1998, *Current Protocols in Molecular Biology*, John Wiley & Sons, NY, which are both incorporated by reference herein in their entireties ), to generate antibodies having a different amino acid sequence, for example to create amino acid substitutions, deletions, and/or insertions.

15           In a specific embodiment, the amino acid sequence of the heavy and/or light chain variable domains may be inspected to identify the sequences of the complementarity determining regions (CDRs) by methods that are well know in the art, e.g., by comparison to known amino acid sequences of other heavy and light chain variable regions to determine the regions of sequence hypervariability. Using routine recombinant DNA techniques, one or more of the CDRs may be inserted within framework regions, e.g., into human framework regions to humanize a non-human antibody, as described supra. The framework regions may be naturally occurring or consensus framework regions, and preferably human framework regions (see, e.g., Chothia et al., *J. Mol. Biol.* 278: 457-479 (1998) for a listing of human framework regions). Preferably, the polynucleotide generated by the combination of the framework regions and CDRs encodes an antibody that specifically binds a polypeptide of the invention. Preferably, as discussed supra, one or more amino acid substitutions may be made within the framework regions, and, preferably, the amino acid substitutions improve binding of the antibody to its antigen. Additionally, such methods may be used to make amino acid substitutions or deletions of one or more variable region cysteine residues participating in an intrachain disulfide bond to generate antibody molecules lacking one or more intrachain disulfide bonds. Other

alterations to the polynucleotide are encompassed by the present invention and within the skill of the art.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., Proc. Natl. Acad. Sci. 81:851-855 (1984); Neuberger et al., Nature 312:604-608 (1984); Takeda et al., Nature 314:452-454 (1985)) by splicing genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. As described supra, a chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region, e.g., humanized antibodies.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778; Bird, Science 242:423- 42 (1988); Huston et al., Proc. Natl. Acad. Sci. USA 85:5879-5883 (1988); and Ward et al., Nature 334:544-54 (1989)) can be adapted to produce single chain antibodies. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide. Techniques for the assembly of functional Fv fragments in E. coli may also be used (Skerra et al., Science 242:1038- 1041 (1988)).

More preferably, a clone encoding an antibody of the present invention may be obtained according to the method described in the Example section herein.

#### Methods of Producing Antibodies

The antibodies of the invention can be produced by any method known in the art for the synthesis of antibodies, in particular, by chemical synthesis or preferably, by recombinant expression techniques.

Recombinant expression of an antibody of the invention, or fragment, derivative or analog thereof, (e.g., a heavy or light chain of an antibody of the invention or a single chain antibody of the invention), requires construction of an expression vector containing a polynucleotide that encodes the antibody. Once a polynucleotide encoding an antibody molecule or a heavy or light chain of an antibody, or portion thereof (preferably containing the heavy or light chain variable domain), of the invention has been obtained, the vector for the production of the

antibody molecule may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing a protein by expressing a polynucleotide containing an antibody encoding nucleotide sequence are described herein. Methods which are well known to those skilled in the art can be used to

5 construct expression vectors containing antibody coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. The invention, thus, provides replicable vectors comprising a nucleotide sequence encoding an antibody molecule of the invention, or a heavy or

10 light chain thereof, or a heavy or light chain variable domain, operably linked to a promoter. Such vectors may include the nucleotide sequence encoding the constant region of the antibody molecule (see, e.g., PCT Publication WO 86/05807; PCT Publication WO 89/01036; and U.S. Patent No. 5,122,464) and the variable domain of the antibody may be cloned into such a vector for expression of the entire heavy or

15 light chain.

The expression vector is transferred to a host cell by conventional techniques and the transfected cells are then cultured by conventional techniques to produce an antibody of the invention. Thus, the invention includes host cells containing a polynucleotide encoding an antibody of the invention, or a heavy or light chain

20 thereof, or a single chain antibody of the invention, operably linked to a heterologous promoter. In preferred embodiments for the expression of double-chained antibodies, vectors encoding both the heavy and light chains may be co-expressed in the host cell<sup>o</sup> for expression of the entire immunoglobulin molecule, as detailed below.

A variety of host-expression vector systems may be utilized to express the

25 antibody molecules of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, express an antibody molecule of the invention in situ. These include but are not limited to microorganisms such as bacteria

30 (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing antibody coding sequences; yeast (e.g., *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors

containing antibody coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing antibody coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing antibody coding sequences; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3 cells) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter). Preferably, bacterial cells such as *Escherichia coli*, and more preferably, eukaryotic cells, especially for the expression of whole recombinant antibody molecule, are used for the expression of a recombinant antibody molecule. For example, mammalian cells such as Chinese hamster ovary cells (CHO), in conjunction with a vector such as the major intermediate early gene promoter element from human cytomegalovirus is an effective expression system for antibodies (Foecking et al., *Gene* 45:101 (1986); Cockett et al., *Bio/Technology* 8:2 (1990)).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the antibody molecule being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions of an antibody molecule, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., *EMBO J.* 2:1791 (1983)), in which the antibody coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, *Nucleic Acids Res.* 13:3101-3109 (1985); Van Heeke & Schuster, *J. Biol. Chem.* 24:5503-5509 (1989)); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption and binding to matrix glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa

protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The antibody coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the antibody coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the antibody molecule in infected hosts. (e.g., see Logan & Shenk, Proc. Natl. Acad. Sci. USA 81:355-359 (1984)). Specific initiation signals may also be required for efficient translation of inserted antibody coding sequences. These signals include the ATG initiation codon and adjacent sequences. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., Methods in Enzymol. 153:51-544 (1987)).

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which

possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERY, BHK, Hela, COS, MDCK, 293, 3T3, WI38, and in particular, breast cancer cell lines such as, for example, BT483, Hs578T, HTB2, BT20 and T47D, and normal mammary gland cell line such as, for example, CRL7030 and Hs578Bst.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express the antibody molecule may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the antibody molecule. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that interact directly or indirectly with the antibody molecule.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler et al., Cell 11:223 (1977)), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, Proc. Natl. Acad. Sci. USA 48:202 (1992)), and adenine phosphoribosyltransferase (Lowy et al., Cell 22:817 (1980)) genes can be employed in tk-, hgp<sup>r</sup>t- or ap<sup>r</sup>t- cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler et al., Natl. Acad. Sci. USA 77:357 (1980); O'Hare et al., Proc. Natl. Acad. Sci. USA 78:1527 (1981)); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, Proc. Natl. Acad. Sci. USA 78:2072 (1981)); neo, which confers resistance to the aminoglycoside G-418 Clinical Pharmacy 12:488-505; Wu and Wu, Biotherapy 3:87-95 (1991); Tolstoshev, Ann. Rev. Pharmacol. Toxicol. 32:573-596 (1993); Mulligan, Science

260:926-932 (1993); and Morgan and Anderson, *Ann. Rev. Biochem.* 62:191-217 (1993); May, 1993, *TIB TECH* 11(5):155-215; and hygromycin (Santerre et al., *Gene* 30:147 (1984)). Methods commonly known in the art of recombinant DNA technology may be routinely applied to select the desired recombinant clone, and such methods are described, for example, in Ausubel et al. (eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons, NY (1993); Kriegler, *Gene Transfer and Expression, A Laboratory Manual*, Stockton Press, NY (1990); and in Chapters 12 and 13, Dracopoli et al. (eds), *Current Protocols in Human Genetics*, John Wiley & Sons, NY (1994); Colberre-Garapin et al., *J. Mol. Biol.* 150:1 (1981), which are incorporated by reference herein in their entireties.

The expression levels of an antibody molecule can be increased by vector amplification (for a review, see Bebbington and Hentschel, *The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells in DNA cloning*, Vol.3. (Academic Press, New York, 1987)). When a marker in the vector system expressing antibody is amplifiable, increase in the level of inhibitor present in culture of host cell will increase the number of copies of the marker gene. Since the amplified region is associated with the antibody gene, production of the antibody will also increase (Crouse et al., *Mol. Cell. Biol.* 3:257 (1983)).

The host cell may be co-transfected with two expression vectors of the invention, the first vector encoding a heavy chain derived polypeptide and the second vector encoding a light chain derived polypeptide. The two vectors may contain identical selectable markers which enable equal expression of heavy and light chain polypeptides. Alternatively, a single vector may be used which encodes, and is capable of expressing, both heavy and light chain polypeptides. In such situations, the light chain should be placed before the heavy chain to avoid an excess of toxic free heavy chain (Proudfoot, *Nature* 322:52 (1986); Kohler, *Proc. Natl. Acad. Sci. USA* 77:2197 (1980)). The coding sequences for the heavy and light chains may comprise cDNA or genomic DNA.

Once an antibody molecule of the invention has been produced by an animal, chemically synthesized, or recombinantly expressed, it may be purified by any method known in the art for purification of an immunoglobulin molecule, for example, by chromatography (e.g., ion exchange, affinity, particularly by affinity for

the specific antigen after Protein A, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins. In addition, the antibodies of the present invention or fragments thereof can be fused to heterologous polypeptide sequences described  
5 herein or otherwise known in the art, to facilitate purification.

The present invention encompasses antibodies recombinantly fused or chemically conjugated (including both covalently and non-covalently conjugations) to a polypeptide (or portion thereof, preferably at least 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 amino acids of the polypeptide) of the present invention to generate fusion  
10 proteins. The fusion does not necessarily need to be direct, but may occur through linker sequences. The antibodies may be specific for antigens other than polypeptides (or portion thereof, preferably at least 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 amino acids of the polypeptide) of the present invention. For example, antibodies may be used to target the polypeptides of the present invention to particular cell types, either  
15 in vitro or in vivo, by fusing or conjugating the polypeptides of the present invention to antibodies specific for particular cell surface receptors. Antibodies fused or conjugated to the polypeptides of the present invention may also be used in in vitro immunoassays and purification methods using methods known in the art. See e.g., Harbor et al., supra, and PCT publication WO 93/21232; EP 439,095; Naramura et al.,  
20 Immunol. Lett. 39:91-99 (1994); U.S. Patent 5,474,981; Gillies et al., PNAS 89:1428-1432 (1992); Fell et al., J. Immunol. 146:2446-2452(1991), which are incorporated by reference in their entirety.

The present invention further includes compositions comprising the polypeptides of the present invention fused or conjugated to antibody domains other  
25 than the variable regions. For example, the polypeptides of the present invention may be fused or conjugated to an antibody Fc region, or portion thereof. The antibody portion fused to a polypeptide of the present invention may comprise the constant region, hinge region, CH1 domain, CH2 domain, and CH3 domain or any combination of whole domains or portions thereof. The polypeptides may also be  
30 fused or conjugated to the above antibody portions to form multimers. For example, Fc portions fused to the polypeptides of the present invention can form dimers through disulfide bonding between the Fc portions. Higher multimeric forms can be

made by fusing the polypeptides to portions of IgA and IgM. Methods for fusing or conjugating the polypeptides of the present invention to antibody portions are known in the art. See, e.g., U.S. Patent Nos. 5,336,603; 5,622,929; 5,359,046; 5,349,053; 5,447,851; 5,112,946; EP 307,434; EP 367,166; PCT publications WO 96/04388; WO 91/06570; Ashkenazi et al., Proc. Natl. Acad. Sci. USA 88:10535-10539 (1991); Zheng et al., J. Immunol. 154:5590-5600 (1995); and Vil et al., Proc. Natl. Acad. Sci. USA 89:11337- 11341(1992) (said references incorporated by reference in their entireties).

As discussed, supra, the polypeptides corresponding to a polypeptide, polypeptide fragment, or a variant of SEQ ID NO:Y may be fused or conjugated to the above antibody portions to increase the in vivo half life of the polypeptides or for use in immunoassays using methods known in the art. Further, the polypeptides corresponding to SEQ ID NO:Y may be fused or conjugated to the above antibody portions to facilitate purification. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP 394,827; Traunecker et al., Nature 331:84-86 (1988). The polypeptides of the present invention fused or conjugated to an antibody having disulfide- linked dimeric structures (due to the IgG) may also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995)). In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP A 232,262). Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, Bennett et al., J. Molecular Recognition 8:52-58 (1995); Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).

Moreover, the antibodies or fragments thereof of the present invention can be fused to marker sequences, such as a peptide to facilitate purification. In preferred

embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Other peptide tags useful for purification include, but are not limited to, the "HA" tag, which corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson et al., Cell 37:767 (1984)) and the "flag" tag.

The present invention further encompasses antibodies or fragments thereof conjugated to a diagnostic or therapeutic agent. The antibodies can be used diagnostically to, for example, monitor the development or progression of a tumor as part of a clinical testing procedure to, e.g., determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, radioactive materials, positron emitting metals using various positron emission tomographies, and nonradioactive paramagnetic metal ions. The detectable substance may be coupled or conjugated either directly to the antibody (or fragment thereof) or indirectly, through an intermediate (such as, for example, a linker known in the art) using techniques known in the art. See, for example, U.S. Patent No. 4,741,900 for metal ions which can be conjugated to antibodies for use as diagnostics according to the present invention. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin; and examples of suitable radioactive material include  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{111}\text{In}$  or  $^{99}\text{Tc}$ .

Further, an antibody or fragment thereof may be conjugated to a therapeutic moiety such as a cytotoxin, e.g., a cytostatic or cytocidal agent, a therapeutic agent or a radioactive metal ion, e.g., alpha-emitters such as, for example,  $^{213}\text{Bi}$ . A cytotoxin

or cytotoxic agent includes any agent that is detrimental to cells. Examples include paclitaxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-  
 5 dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologues thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU),  
 10 cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

15 The conjugates of the invention can be used for modifying a given biological response, the therapeutic agent or drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a  
 20 protein such as tumor necrosis factor,  $\alpha$ -interferon,  $\beta$ -interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator, an apoptotic agent, e.g., TNF- $\alpha$ , TNF- $\beta$ , AIM I (See, International Publication No. WO 97/33899), AIM II (See, International Publication No. WO 97/34911), Fas Ligand (Takahashi et al., Int. Immunol., 6:1567-1574 (1994)), VEGI (See, International Publication No.  
 25 WO 99/23105), a thrombotic agent or an anti-angiogenic agent, e.g., angiostatin or endostatin; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

30 Antibodies may also be attached to solid supports, which are particularly useful for immunoassays or purification of the target antigen. Such solid supports

include, but are not limited to, glass, cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, e.g., Arnon et al., "Monoclonal Antibodies For Immunotargeting Of  
5 Drugs In Cancer Therapy", in *Monoclonal Antibodies And Cancer Therapy*, Reisfeld et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., "Antibodies For Drug Delivery", in *Controlled Drug Delivery* (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in *Monoclonal Antibodies '84: Biological And  
10 Clinical Applications*, Pinchera et al. (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in *Monoclonal Antibodies For Cancer Detection And Therapy*, Baldwin et al. (eds.), pp. 303-16 (Academic Press 1985), and Thorpe et al., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", *Immunol. Rev.* 62:119-58  
15 (1982).

Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980, which is incorporated herein by reference in its entirety.

An antibody, with or without a therapeutic moiety conjugated to it,  
20 administered alone or in combination with cytotoxic factor(s) and/or cytokine(s) can be used as a therapeutic.

The present invention also encompasses the creation of synthetic antibodies directed against the polypeptides of the present invention. One example of synthetic antibodies is described in Radrizzani, M., et al., *Medicina*, (Aires), 59(6):753-8,  
25 (1999)). Recently, a new class of synthetic antibodies has been described and are referred to as molecularly imprinted polymers (MIPs) (Semorex, Inc.). Antibodies, peptides, and enzymes are often used as molecular recognition elements in chemical and biological sensors. However, their lack of stability and signal transduction mechanisms limits their use as sensing devices. Molecularly imprinted polymers  
30 (MIPs) are capable of mimicking the function of biological receptors but with less stability constraints. Such polymers provide high sensitivity and selectivity while maintaining excellent thermal and mechanical stability. MIPs have the ability to bind

to small molecules and to target molecules such as organics and proteins' with equal or greater potency than that of natural antibodies. These "super" MIPs have higher affinities for their target and thus require lower concentrations for efficacious binding.

During synthesis, the MIPs are imprinted so as to have complementary size,  
5 shape, charge and functional groups of the selected target by using the target molecule itself (such as a polypeptide, antibody, etc.), or a substance having a very similar structure, as its "print" or "template." MIPs can be derivatized with the same reagents afforded to antibodies. For example, fluorescent 'super' MIPs can be coated onto beads or wells for use in highly sensitive separations or assays, or for use in high  
10 throughput screening of proteins.

Moreover, MIPs based upon the structure of the polypeptide(s) of the present invention may be useful in screening for compounds that bind to the polypeptide(s) of the invention. Such a MIP would serve the role of a synthetic "receptor" by minimicking the native architecture of the polypeptide. In fact, the ability of a MIP to  
15 serve the role of a synthetic receptor has already been demonstrated for the estrogen receptor (Ye, L., Yu, Y., Mosbach, K, *Analyst.*, 126(6):760-5, (2001); Dickert, F, L., Hayden, O., Halikias, K, P, *Analyst.*, 126(6):766-71, (2001)). A synthetic receptor may either be mimicked in its entirety (e.g., as the entire protein), or mimicked as a series of short peptides corresponding to the protein (Rachkov, A., Minoura, N,  
20 *Biochim, Biophys, Acta.*, 1544(1-2):255-66, (2001)). Such a synthetic receptor MIPs may be employed in any one or more of the screening methods described elsewhere herein.

MIPs have also been shown to be useful in "sensing" the presence of its mimicked molecule (Cheng, Z., Wang, E., Yang, X, *Biosens, Bioelectron.*, 16(3):179-  
25 85, (2001) ; Jenkins, A, L., Yin, R., Jensen, J. L, *Analyst.*, 126(6):798-802, (2001) ; Jenkins, A, L., Yin, R., Jensen, J. L, *Analyst.*, 126(6):798-802, (2001)). For example, a MIP designed using a polypeptide of the present invention may be used in assays designed to identify, and potentially quantitate, the level of said polypeptide in a sample. Such a MIP may be used as a substitute for any component described in the  
30 assays, or kits, provided herein (e.g., ELISA, etc.).

A number of methods may be employed to create MIPs to a specific receptor, ligand, polypeptide, peptide, organic molecule. Several preferred methods are

described by Esteban et al in J. Anal, Chem., 370(7):795-802, (2001), which is hereby incorporated herein by reference in its entirety in addition to any references cited therein. Additional methods are known in the art and are encompassed by the present invention, such as for example, Hart, B, R., Shea, K, J. J. Am. Chem, Soc.,  
5 123(9):2072-3, (2001); and Quaglia, M., Chenon, K., Hall, A, J., De, Lorenzi, E., Sellergren, B, J. Am. Chem, Soc., 123(10):2146-54, (2001); which are hereby incorporated by reference in their entirety herein.

#### Uses for Antibodies directed against polypeptides of the invention

The antibodies of the present invention have various utilities. For example,  
10 such antibodies may be used in diagnostic assays to detect the presence or quantification of the polypeptides of the invention in a sample. Such a diagnostic assay may be comprised of at least two steps. The first, subjecting a sample with the antibody, wherein the sample is a tissue (e.g., human, animal, etc.), biological fluid (e.g., blood, urine, sputum, semen, amniotic fluid, saliva, etc.), biological extract (e.g.,  
15 tissue or cellular homogenate, etc.), a protein microchip (e.g., See Arenkov P, et al., Anal Biochem., 278(2):123-131 (2000)), or a chromatography column, etc. And a second step involving the quantification of antibody bound to the substrate. Alternatively, the method may additionally involve a first step of attaching the antibody, either covalently, electrostatically, or reversibly, to a solid support, and a  
20 second step of subjecting the bound antibody to the sample, as defined above and elsewhere herein.

Various diagnostic assay techniques are known in the art, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogenous phases (Zola, Monoclonal  
25 Antibodies: A Manual of Techniques, CRC Press, Inc., (1987), pp147-158). The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as  $^2\text{H}$ ,  $^{14}\text{C}$ ,  $^{32}\text{P}$ , or  $^{125}\text{I}$ , a florescent or chemiluminescent compound, such as  
30 fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase, green fluorescent protein, or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety

may be employed, including those methods described by Hunter et al., *Nature*, 144:945 (1962); Dafvid et al., *Biochem.*, 13:1014 (1974); Pain et al., *J. Immunol. Metho.*, 40:219(1981); and Nygren, J. *Histochem. And Cytochem.*, 30:407 (1982).

Antibodies directed against the polypeptides of the present invention are  
5 useful for the affinity purification of such polypeptides from recombinant cell culture or natural sources. In this process, the antibodies against a particular polypeptide are immobilized on a suitable support, such as a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the polypeptides to be purified, and thereafter the support is  
10 washed with a suitable solvent that will remove substantially all the material in the sample except for the desired polypeptides, which are bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the desired polypeptide from the antibody.

#### Immunophenotyping

15 The antibodies of the invention may be utilized for immunophenotyping of cell lines and biological samples. The translation product of the gene of the present invention may be useful as a cell specific marker, or more specifically as a cellular marker that is differentially expressed at various stages of differentiation and/or maturation of particular cell types. Monoclonal antibodies directed against a specific  
20 epitope, or combination of epitopes, will allow for the screening of cellular populations expressing the marker. Various techniques can be utilized using monoclonal antibodies to screen for cellular populations expressing the marker(s), and include magnetic separation using antibody-coated magnetic beads, "panning" with antibody attached to a solid matrix (i.e., plate), and flow cytometry (See, e.g., U.S.  
25 Patent 5,985,660; and Morrison et al., *Cell*, 96:737-49 (1999)).

These techniques allow for the screening of particular populations of cells, such as might be found with hematological malignancies (i.e. minimal residual disease (MRD) in acute leukemic patients) and "non-self" cells in transplantations to prevent Graft-versus-Host Disease (GVHD). Alternatively, these techniques allow for  
30 the screening of hematopoietic stem and progenitor cells capable of undergoing proliferation and/or differentiation, as might be found in human umbilical cord blood.

## Assays For Antibody Binding

The antibodies of the invention may be assayed for immunospecific binding by any method known in the art. The immunoassays which can be used include but are not limited to competitive and non-competitive assay systems using techniques such as western blots, radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoprecipitation assays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays, fluorescent immunoassays, protein A immunoassays, to name but a few. Such assays are routine and well known in the art (see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York, which is incorporated by reference herein in its entirety). Exemplary immunoassays are described briefly below (but are not intended by way of limitation).

Immunoprecipitation protocols generally comprise lysing a population of cells in a lysis buffer such as RIPA buffer (1% NP-40 or Triton X- 100, 1% sodium deoxycholate, 0.1% SDS, 0.15 M NaCl, 0.01 M sodium phosphate at pH 7.2, 1% Trasylol) supplemented with protein phosphatase and/or protease inhibitors (e.g., EDTA, PMSF, aprotinin, sodium vanadate), adding the antibody of interest to the cell lysate, incubating for a period of time (e.g., 1-4 hours) at 4° C, adding protein A and/or protein G sepharose beads to the cell lysate, incubating for about an hour or more at 4° C, washing the beads in lysis buffer and resuspending the beads in SDS/sample buffer. The ability of the antibody of interest to immunoprecipitate a particular antigen can be assessed by, e.g., western blot analysis. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the binding of the antibody to an antigen and decrease the background (e.g., pre-clearing the cell lysate with sepharose beads). For further discussion regarding immunoprecipitation protocols see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 10.16.1.

Western blot analysis generally comprises preparing protein samples, electrophoresis of the protein samples in a polyacrylamide gel (e.g., 8%- 20% SDS-PAGE depending on the molecular weight of the antigen), transferring the protein sample from the polyacrylamide gel to a membrane such as nitrocellulose, PVDF or

nylon, blocking the membrane in blocking solution (e.g., PBS with 3% BSA or non-fat milk), washing the membrane in washing buffer (e.g., PBS-Tween 20), blocking the membrane with primary antibody (the antibody of interest) diluted in blocking buffer, washing the membrane in washing buffer, blocking the membrane with a  
5 secondary antibody (which recognizes the primary antibody, e.g., an anti-human antibody) conjugated to an enzymatic substrate (e.g., horseradish peroxidase or alkaline phosphatase) or radioactive molecule (e.g.,  $^{32}\text{P}$  or  $^{125}\text{I}$ ) diluted in blocking buffer, washing the membrane in wash buffer, and detecting the presence of the antigen. One of skill in the art would be knowledgeable as to the parameters that can  
10 be modified to increase the signal detected and to reduce the background noise. For further discussion regarding western blot protocols see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 10.8.1.

ELISAs comprise preparing antigen, coating the well of a 96 well microtiter  
15 plate with the antigen, adding the antibody of interest conjugated to a detectable compound such as an enzymatic substrate (e.g., horseradish peroxidase or alkaline phosphatase) to the well and incubating for a period of time, and detecting the presence of the antigen. In ELISAs the antibody of interest does not have to be conjugated to a detectable compound; instead, a second antibody (which recognizes  
20 the antibody of interest) conjugated to a detectable compound may be added to the well. Further, instead of coating the well with the antigen, the antibody may be coated to the well. In this case, a second antibody conjugated to a detectable compound may be added following the addition of the antigen of interest to the coated well. One of skill in the art would be knowledgeable as to the parameters that can be modified to  
25 increase the signal detected as well as other variations of ELISAs known in the art. For further discussion regarding ELISAs see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 11.2.1.

The binding affinity of an antibody to an antigen and the off-rate of an  
30 antibody-antigen interaction can be determined by competitive binding assays. One example of a competitive binding assay is a radioimmunoassay comprising the incubation of labeled antigen (e.g.,  $^3\text{H}$  or  $^{125}\text{I}$ ) with the antibody of interest in the

presence of increasing amounts of unlabeled antigen, and the detection of the antibody bound to the labeled antigen. The affinity of the antibody of interest for a particular antigen and the binding off-rates can be determined from the data by scatchard plot analysis. Competition with a second antibody can also be determined  
5 using radioimmunoassays. In this case, the antigen is incubated with antibody of interest conjugated to a labeled compound (e.g.,  $^3\text{H}$  or  $^{125}\text{I}$ ) in the presence of increasing amounts of an unlabeled second antibody.

#### Therapeutic Uses Of Antibodies

The present invention is further directed to antibody-based therapies which  
10 involve administering antibodies of the invention to an animal, preferably a mammal, and most preferably a human, patient for treating one or more of the disclosed diseases, disorders, or conditions. Therapeutic compounds of the invention include, but are not limited to, antibodies of the invention (including fragments, analogs and derivatives thereof as described herein) and nucleic acids encoding antibodies of the  
15 invention (including fragments, analogs and derivatives thereof and anti-idiotypic antibodies as described herein). The antibodies of the invention can be used to treat, inhibit or prevent diseases, disorders or conditions associated with aberrant expression and/or activity of a polypeptide of the invention, including, but not limited to, any one or more of the diseases, disorders, or conditions described herein. The treatment  
20 and/or prevention of diseases, disorders, or conditions associated with aberrant expression and/or activity of a polypeptide of the invention includes, but is not limited to, alleviating symptoms associated with those diseases, disorders or conditions. Antibodies of the invention may be provided in pharmaceutically acceptable compositions as known in the art or as described herein.

25 A summary of the ways in which the antibodies of the present invention may be used therapeutically includes binding polynucleotides or polypeptides of the present invention locally or systemically in the body or by direct cytotoxicity of the antibody, e.g. as mediated by complement (CDC) or by effector cells (ADCC). Some of these approaches are described in more detail below. Armed with the teachings  
30 provided herein, one of ordinary skill in the art will know how to use the antibodies of the present invention for diagnostic, monitoring or therapeutic purposes without undue experimentation.

The antibodies of this invention may be advantageously utilized in combination with other monoclonal or chimeric antibodies, or with lymphokines or hematopoietic growth factors (such as, e.g., IL-2, IL-3 and IL-7), for example, which serve to increase the number or activity of effector cells which interact with the antibodies.

The antibodies of the invention may be administered alone or in combination with other types of treatments (e.g., radiation therapy, chemotherapy, hormonal therapy, immunotherapy and anti-tumor agents). Generally, administration of products of a species origin or species reactivity (in the case of antibodies) that is the same species as that of the patient is preferred. Thus, in a preferred embodiment, human antibodies, fragments derivatives, analogs, or nucleic acids, are administered to a human patient for therapy or prophylaxis.

It is preferred to use high affinity and/or potent in vivo inhibiting and/or neutralizing antibodies against polypeptides or polynucleotides of the present invention, fragments or regions thereof, for both immunoassays directed to and therapy of disorders related to polynucleotides or polypeptides, including fragments thereof, of the present invention. Such antibodies, fragments, or regions, will preferably have an affinity for polynucleotides or polypeptides of the invention, including fragments thereof. Preferred binding affinities include those with a dissociation constant or  $K_d$  less than  $5 \times 10^{-2}$  M,  $10^{-2}$  M,  $5 \times 10^{-3}$  M,  $10^{-3}$  M,  $5 \times 10^{-4}$  M,  $10^{-4}$  M,  $5 \times 10^{-5}$  M,  $10^{-5}$  M,  $5 \times 10^{-6}$  M,  $10^{-6}$  M,  $5 \times 10^{-7}$  M,  $10^{-7}$  M,  $5 \times 10^{-8}$  M,  $10^{-8}$  M,  $5 \times 10^{-9}$  M,  $10^{-9}$  M,  $5 \times 10^{-10}$  M,  $10^{-10}$  M,  $5 \times 10^{-11}$  M,  $10^{-11}$  M,  $5 \times 10^{-12}$  M,  $10^{-12}$  M,  $5 \times 10^{-13}$  M,  $10^{-13}$  M,  $5 \times 10^{-14}$  M,  $10^{-14}$  M,  $5 \times 10^{-15}$  M, and  $10^{-15}$  M.

Antibodies directed against polypeptides of the present invention are useful for inhibiting allergic reactions in animals. For example, by administering a therapeutically acceptable dose of an antibody, or antibodies, of the present invention, or a cocktail of the present antibodies, or in combination with other antibodies of varying sources, the animal may not elicit an allergic response to antigens.

Likewise, one could envision cloning the gene encoding an antibody directed against a polypeptide of the present invention, said polypeptide having the potential to elicit an allergic and/or immune response in an organism, and transforming the

organism with said antibody gene such that it is expressed (e.g., constitutively, inducibly, etc.) in the organism. Thus, the organism would effectively become resistant to an allergic response resulting from the ingestion or presence of such an immune/allergic reactive polypeptide. Moreover, such a use of the antibodies of the present invention may have particular utility in preventing and/or ameliorating autoimmune diseases and/or disorders, as such conditions are typically a result of antibodies being directed against endogenous proteins. For example, in the instance where the polypeptide of the present invention is responsible for modulating the immune response to auto-antigens, transforming the organism and/or individual with a construct comprising any of the promoters disclosed herein or otherwise known in the art, in addition, to a polynucleotide encoding the antibody directed against the polypeptide of the present invention could effectively inhibit the organism's immune system from eliciting an immune response to the auto-antigen(s). Detailed descriptions of therapeutic and/or gene therapy applications of the present invention are provided elsewhere herein.

Alternatively, antibodies of the present invention could be produced in a plant (e.g., cloning the gene of the antibody directed against a polypeptide of the present invention, and transforming a plant with a suitable vector comprising said gene for constitutive expression of the antibody within the plant), and the plant subsequently ingested by an animal, thereby conferring temporary immunity to the animal for the specific antigen the antibody is directed towards (See, for example, US Patent Nos. 5,914,123 and 6,034,298).

In another embodiment, antibodies of the present invention, preferably polyclonal antibodies, more preferably monoclonal antibodies, and most preferably single-chain antibodies, can be used as a means of inhibiting gene expression of a particular gene, or genes, in a human, mammal, and/or other organism. See, for example, International Publication Number WO 00/05391, published 2/3/00, to Dow Agrosciences LLC. The application of such methods for the antibodies of the present invention are known in the art, and are more particularly described elsewhere herein.

In yet another embodiment, antibodies of the present invention may be useful for multimerizing the polypeptides of the present invention. For example, certain proteins may confer enhanced biological activity when present in a multimeric state

(i.e., such enhanced activity may be due to the increased effective concentration of such proteins whereby more protein is available in a localized location).

#### Antibody-based Gene Therapy

In a specific embodiment, nucleic acids comprising sequences encoding  
5 antibodies or functional derivatives thereof, are administered to treat, inhibit or prevent a disease or disorder associated with aberrant expression and/or activity of a polypeptide of the invention, by way of gene therapy. Gene therapy refers to therapy performed by the administration to a subject of an expressed or expressible nucleic acid. In this embodiment of the invention, the nucleic acids produce their encoded  
10 protein that mediates a therapeutic effect.

Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

For general reviews of the methods of gene therapy, see Goldspiel et al., Clinical Pharmacy 12:488-505 (1993); Wu and Wu, Biotherapy 3:87-95 (1991);  
15 Tolstoshev, Ann. Rev. Pharmacol. Toxicol. 32:573-596 (1993); Mulligan, Science 260:926-932 (1993); and Morgan and Anderson, Ann. Rev. Biochem. 62:191-217 (1993); May, TIBTECH 11(5):155-215 (1993). Methods commonly known in the art of recombinant DNA technology which can be used are described in Ausubel et al. (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, NY (1993); and  
20 Kriegler, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY (1990).

In a preferred aspect, the compound comprises nucleic acid sequences encoding an antibody, said nucleic acid sequences being part of expression vectors that express the antibody or fragments or chimeric proteins or heavy or light chains  
25 thereof in a suitable host. In particular, such nucleic acid sequences have promoters operably linked to the antibody coding region, said promoter being inducible or constitutive, and, optionally, tissue- specific. In another particular embodiment, nucleic acid molecules are used in which the antibody coding sequences and any other desired sequences are flanked by regions that promote homologous recombination at a  
30 desired site in the genome, thus providing for intrachromosomal expression of the antibody encoding nucleic acids (Koller and Smithies, Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); Zijlstra et al., Nature 342:435-438 (1989). In specific

embodiments, the expressed antibody molecule is a single chain antibody; alternatively, the nucleic acid sequences include sequences encoding both the heavy and light chains, or fragments thereof, of the antibody.

5 Delivery of the nucleic acids into a patient may be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid- carrying vectors, or indirect, in which case, cells are first transformed with the nucleic acids in vitro, then transplanted into the patient. These two approaches are known, respectively, as in vivo or ex vivo gene therapy.

10 In a specific embodiment, the nucleic acid sequences are directly administered in vivo, where it is expressed to produce the encoded product. This can be accomplished by any of numerous methods known in the art, e.g., by constructing them as part of an appropriate nucleic acid expression vector and administering it so that they become intracellular, e.g., by infection using defective or attenuated retrovirals or other viral vectors (see U.S. Patent No. 4,980,286), or by direct injection  
15 of naked DNA, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in liposomes, microparticles, or microcapsules, or by administering them in linkage to a peptide which is known to enter the nucleus, by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see, e.g., Wu and Wu,  
20 J. Biol. Chem. 262:4429-4432 (1987)) (which can be used to target cell types specifically expressing the receptors), etc. In another embodiment, nucleic acid-ligand complexes can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted in vivo for cell specific uptake  
25 and expression, by targeting a specific receptor (see, e.g., PCT Publications WO 92/06180; WO 92/22635; WO92/20316; WO93/14188, WO 93/20221). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination (Koller and Smithies, Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); Zijlstra et al., Nature 342:435-438 (1989)).

30 In a specific embodiment, viral vectors that contains nucleic acid sequences encoding an antibody of the invention are used. For example, a retroviral vector can be used (see Miller et al., Meth. Enzymol. 217:581-599 (1993)). These retroviral

vectors contain the components necessary for the correct packaging of the viral genome and integration into the host cell DNA. The nucleic acid sequences encoding the antibody to be used in gene therapy are cloned into one or more vectors, which facilitates delivery of the gene into a patient. More detail about retroviral vectors can be found in Boesen et al., *Biotherapy* 6:291-302 (1994), which describes the use of a retroviral vector to deliver the *mdr1* gene to hematopoietic stem cells in order to make the stem cells more resistant to chemotherapy. Other references illustrating the use of retroviral vectors in gene therapy are: Clowes et al., *J. Clin. Invest.* 93:644-651 (1994); Kiem et al., *Blood* 83:1467-1473 (1994); Salmons and Gunzberg, *Human Gene Therapy* 4:129-141 (1993); and Grossman and Wilson, *Curr. Opin. in Genetics and Devel.* 3:110-114 (1993).

Adenoviruses are other viral vectors that can be used in gene therapy. Adenoviruses are especially attractive vehicles for delivering genes to respiratory epithelia. Adenoviruses naturally infect respiratory epithelia where they cause a mild disease. Other targets for adenovirus-based delivery systems are liver, the central nervous system, endothelial cells, and muscle. Adenoviruses have the advantage of being capable of infecting non-dividing cells. Kozarsky and Wilson, *Current Opinion in Genetics and Development* 3:499-503 (1993) present a review of adenovirus-based gene therapy. Bout et al., *Human Gene Therapy* 5:3-10 (1994) demonstrated the use of adenovirus vectors to transfer genes to the respiratory epithelia of rhesus monkeys. Other instances of the use of adenoviruses in gene therapy can be found in Rosenfeld et al., *Science* 252:431-434 (1991); Rosenfeld et al., *Cell* 68:143-155 (1992); Mastrangeli et al., *J. Clin. Invest.* 91:225-234 (1993); PCT Publication WO94/12649; and Wang, et al., *Gene Therapy* 2:775-783 (1995). In a preferred embodiment, adenovirus vectors are used.

Adeno-associated virus (AAV) has also been proposed for use in gene therapy (Walsh et al., *Proc. Soc. Exp. Biol. Med.* 204:289-300 (1993); U.S. Patent No. 5,436,146).

Another approach to gene therapy involves transferring a gene to cells in tissue culture by such methods as electroporation, lipofection, calcium phosphate mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to

isolate those cells that have taken up and are expressing the transferred gene. Those cells are then delivered to a patient.

In this embodiment, the nucleic acid is introduced into a cell prior to administration in vivo of the resulting recombinant cell. Such introduction can be carried out by any method known in the art, including but not limited to transfection, electroporation, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells (see, e.g., Loeffler and Behr, *Meth. Enzymol.* 217:599-618 (1993); Cohen et al., *Meth. Enzymol.* 217:618-644 (1993); Cline, *Pharmac. Ther.* 29:69-92m (1985) and may be used in accordance with the present invention, provided that the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably heritable and expressible by its cell progeny.

The resulting recombinant cells can be delivered to a patient by various methods known in the art. Recombinant blood cells (e.g., hematopoietic stem or progenitor cells) are preferably administered intravenously. The amount of cells envisioned for use depends on the desired effect, patient state, etc., and can be determined by one skilled in the art.

Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, available cell type, and include but are not limited to epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes; blood cells such as Tlymphocytes, Blymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, granulocytes; various stem or progenitor cells, in particular hematopoietic stem or progenitor cells, e.g., as obtained from bone marrow, umbilical cord blood, peripheral blood, fetal liver, etc.

In a preferred embodiment, the cell used for gene therapy is autologous to the patient.

In an embodiment in which recombinant cells are used in gene therapy, nucleic acid sequences encoding an antibody are introduced into the cells such that

they are expressible by the cells or their progeny, and the recombinant cells are then administered in vivo for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem and/or progenitor cells which can be isolated and maintained in vitro can potentially be used in accordance with this embodiment of the present invention (see e.g. PCT Publication WO 94/08598; Stemple and Anderson, Cell 71:973-985 (1992); Rheinwald, Meth. Cell Bio. 21A:229 (1980); and Pittelkow and Scott, Mayo Clinic Proc. 61:771 (1986)).

In a specific embodiment, the nucleic acid to be introduced for purposes of gene therapy comprises an inducible promoter operably linked to the coding region, such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription. Demonstration of Therapeutic or Prophylactic Activity

The compounds or pharmaceutical compositions of the invention are preferably tested in vitro, and then in vivo for the desired therapeutic or prophylactic activity, prior to use in humans. For example, in vitro assays to demonstrate the therapeutic or prophylactic utility of a compound or pharmaceutical composition include, the effect of a compound on a cell line or a patient tissue sample. The effect of the compound or composition on the cell line and/or tissue sample can be determined utilizing techniques known to those of skill in the art including, but not limited to, rosette formation assays and cell lysis assays. In accordance with the invention, in vitro assays which can be used to determine whether administration of a specific compound is indicated, include in vitro cell culture assays in which a patient tissue sample is grown in culture, and exposed to or otherwise administered a compound, and the effect of such compound upon the tissue sample is observed.

#### Therapeutic/Prophylactic Administration and Compositions

The invention provides methods of treatment, inhibition and prophylaxis by administration to a subject of an effective amount of a compound or pharmaceutical composition of the invention, preferably an antibody of the invention. In a preferred aspect, the compound is substantially purified (e.g., substantially free from substances that limit its effect or produce undesired side-effects). The subject is preferably an animal, including but not limited to animals such as cows, pigs, horses, chickens, cats, dogs, etc., and is preferably a mammal, and most preferably human.

Formulations and methods of administration that can be employed when the compound comprises a nucleic acid or an immunoglobulin are described above; additional appropriate formulations and routes of administration can be selected from among those described herein below.

5           Various delivery systems are known and can be used to administer a compound of the invention, e.g., encapsulation in liposomes, microparticles, microcapsules, recombinant cells capable of expressing the compound, receptor-mediated endocytosis (see, e.g., Wu and Wu, J. Biol. Chem. 262:4429-4432 (1987)), construction of a nucleic acid as part of a retroviral or other vector, etc. Methods of  
10 introduction include but are not limited to intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral routes. The compounds or compositions may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (e.g., oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with  
15 other biologically active agents. Administration can be systemic or local. In addition, it may be desirable to introduce the pharmaceutical compounds or compositions of the invention into the central nervous system by any suitable route, including intraventricular and intrathecal injection; intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir, such as an  
20 Ommaya reservoir. Pulmonary administration can also be employed, e.g., by use of an inhaler or nebulizer, and formulation with an aerosolizing agent.

          In a specific embodiment, it may be desirable to administer the pharmaceutical compounds or compositions of the invention locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion  
25 during surgery, topical application, e.g., in conjunction with a wound dressing after surgery, by injection, by means of a catheter, by means of a suppository, or by means of an implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers. Preferably, when administering a protein, including an antibody, of the invention, care must be taken to  
30 use materials to which the protein does not absorb.

          In another embodiment, the compound or composition can be delivered in a vesicle, in particular a liposome (see Langer, Science 249:1527-1533 (1990); Treat et

al., in *Liposomes in the Therapy of Infectious Disease and Cancer*, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 353- 365 (1989); Lopez-Berestein, *ibid.*, pp. 317-327; see generally *ibid.*)

5 In yet another embodiment, the compound or composition can be delivered in a controlled release system. In one embodiment, a pump may be used (see Langer, *supra*; Sefton, *CRC Crit. Ref. Biomed. Eng.* 14:201 (1987); Buchwald et al., *Surgery* 88:507 (1980); Saudek et al., *N. Engl. J. Med.* 321:574 (1989)). In another embodiment, polymeric materials can be used (see *Medical Applications of Controlled Release*, Langer and Wise (eds.), CRC Pres., Boca Raton, Florida (1974);  
10 *Controlled Drug Bioavailability, Drug Product Design and Performance*, Smolen and Ball (eds.), Wiley, New York (1984); Ranger and Peppas, J., *Macromol. Sci. Rev. Macromol. Chem.* 23:61 (1983); see also Levy et al., *Science* 228:190 (1985); During et al., *Ann. Neurol.* 25:351 (1989); Howard et al., *J. Neurosurg.* 71:105 (1989)). In yet another embodiment, a controlled release system can be placed in proximity of the  
15 therapeutic target, i.e., the brain, thus requiring only a fraction of the systemic dose (see, e.g., Goodson, in *Medical Applications of Controlled Release*, *supra*, vol. 2, pp. 115-138 (1984)).

Other controlled release systems are discussed in the review by Langer (*Science* 249:1527-1533 (1990)).

20 In a specific embodiment where the compound of the invention is a nucleic acid encoding a protein, the nucleic acid can be administered in vivo to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by use of a retroviral vector (see U.S. Patent No. 4,980,286), or by direct injection, or by  
25 use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox-like peptide which is known to enter the nucleus (see e.g., Joliot et al., *Proc. Natl. Acad. Sci. USA* 88:1864-1868 (1991)), etc. Alternatively, a nucleic acid can be introduced intracellularly and incorporated within host cell DNA  
30 for expression, by homologous recombination.

The present invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically effective amount of a compound, and a

pharmaceutically acceptable carrier. In a specific embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered. Such pharmaceutical carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered intravenously. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. These compositions can take the form of solutions, suspensions, emulsion, tablets, pills, capsules, powders, sustained-release formulations and the like. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Examples of suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E.W. Martin. Such compositions will contain a therapeutically effective amount of the compound, preferably in purified form, together with a suitable amount of carrier so as to provide the form for proper administration to the patient. The formulation should suit the mode of administration.

In a preferred embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition may also include a solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients are

supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion  
5 bottle containing sterile pharmaceutical grade water or saline. Where the composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

The compounds of the invention can be formulated as neutral or salt forms. Pharmaceutically acceptable salts include those formed with anions such as those  
10 derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., and those formed with cations such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc.

The amount of the compound of the invention which will be effective in the  
15 treatment, inhibition and prevention of a disease or disorder associated with aberrant expression and/or activity of a polypeptide of the invention can be determined by standard clinical techniques. In addition, in vitro assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the  
20 disease or disorder, and should be decided according to the judgment of the practitioner and each patient's circumstances. Effective doses may be extrapolated from dose-response curves derived from in vitro or animal model test systems.

For antibodies, the dosage administered to a patient is typically 0.1 mg/kg to 100 mg/kg of the patient's body weight. Preferably, the dosage administered to a  
25 patient is between 0.1 mg/kg and 20 mg/kg of the patient's body weight, more preferably 1 mg/kg to 10 mg/kg of the patient's body weight. Generally, human antibodies have a longer half-life within the human body than antibodies from other species due to the immune response to the foreign polypeptides. Thus, lower dosages of human antibodies and less frequent administration is often possible. Further, the  
30 dosage and frequency of administration of antibodies of the invention may be reduced by enhancing uptake and tissue penetration (e.g., into the brain) of the antibodies by modifications such as, for example, lipidation.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration...

#### Diagnosis and Imaging With Antibodies

Labeled antibodies, and derivatives and analogs thereof, which specifically bind to a polypeptide of interest can be used for diagnostic purposes to detect, diagnose, or monitor diseases, disorders, and/or conditions associated with the aberrant expression and/or activity of a polypeptide of the invention. The invention provides for the detection of aberrant expression of a polypeptide of interest, comprising (a) assaying the expression of the polypeptide of interest in cells or body fluid of an individual using one or more antibodies specific to the polypeptide interest and (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of aberrant expression.

The invention provides a diagnostic assay for diagnosing a disorder, comprising (a) assaying the expression of the polypeptide of interest in cells or body fluid of an individual using one or more antibodies specific to the polypeptide interest and (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a particular disorder. With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Antibodies of the invention can be used to assay protein levels in a biological sample using classical immunohistological methods known to those of skill in the art (e.g., see Jalkanen, et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, et al., J. Cell .

Biol. 105:3087-3096 (1987)). Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase; 5 radioisotopes, such as iodine ( $^{125}\text{I}$ ,  $^{121}\text{I}$ ), carbon ( $^{14}\text{C}$ ), sulfur ( $^{35}\text{S}$ ), tritium ( $^3\text{H}$ ), indium ( $^{112}\text{In}$ ), and technetium ( $^{99}\text{Tc}$ ); luminescent labels, such as luminol; and fluorescent labels, such as fluorescein and rhodamine, and biotin.

One aspect of the invention is the detection and diagnosis of a disease or disorder associated with aberrant expression of a polypeptide of interest in an animal, 10 preferably a mammal and most preferably a human. In one embodiment, diagnosis comprises: a) administering (for example, parenterally, subcutaneously, or intraperitoneally) to a subject an effective amount of a labeled molecule which specifically binds to the polypeptide of interest; b) waiting for a time interval following the administering for permitting the labeled molecule to preferentially 15 concentrate at sites in the subject where the polypeptide is expressed (and for unbound labeled molecule to be cleared to background level); c) determining background level; and d) detecting the labeled molecule in the subject, such that detection of labeled molecule above the background level indicates that the subject has a particular disease or disorder associated with aberrant expression of the 20 polypeptide of interest. Background level can be determined by various methods including, comparing the amount of labeled molecule detected to a standard value previously determined for a particular system.

It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce 25 diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of  $^{99\text{mTc}}$ . The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of 30 Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).

Depending on several variables, including the type of label used and the mode of administration, the time interval following the administration for permitting the labeled molecule to preferentially concentrate at sites in the subject and for unbound labeled molecule to be cleared to background level is 6 to 48 hours or 6 to 24 hours or  
5 6 to 12 hours. In another embodiment the time interval following administration is 5 to 20 days or 5 to 10 days.

In an embodiment, monitoring of the disease or disorder is carried out by repeating the method for diagnosing the disease or disease, for example, one month after initial diagnosis, six months after initial diagnosis, one year after initial  
10 diagnosis, etc.

Presence of the labeled molecule can be detected in the patient using methods known in the art for in vivo scanning. These methods depend upon the type of label used. Skilled artisans will be able to determine the appropriate method for detecting a particular label. Methods and devices that may be used in the diagnostic methods of  
15 the invention include, but are not limited to, computed tomography (CT), whole body scan such as position emission tomography (PET), magnetic resonance imaging (MRI), and sonography.

In a specific embodiment, the molecule is labeled with a radioisotope and is detected in the patient using a radiation responsive surgical instrument (Thurston et al., U.S. Patent No. 5,441,050). In another embodiment, the molecule is labeled with a  
20 fluorescent compound and is detected in the patient using a fluorescence responsive scanning instrument. In another embodiment, the molecule is labeled with a positron emitting metal and is detected in the patient using positron emission-tomography. In yet another embodiment, the molecule is labeled with a paramagnetic label and is  
25 detected in a patient using magnetic resonance imaging (MRI).

#### Kits

The present invention provides kits that can be used in the above methods. In one embodiment, a kit comprises an antibody of the invention, preferably a purified antibody, in one or more containers. In a specific embodiment, the kits of the present  
30 invention contain a substantially isolated polypeptide comprising an epitope which is specifically immunoreactive with an antibody included in the kit. Preferably, the kits of the present invention further comprise a control antibody which does not react with

the polypeptide of interest. In another specific embodiment, the kits of the present invention contain a means for detecting the binding of an antibody to a polypeptide of interest (e.g., the antibody may be conjugated to a detectable substrate such as a fluorescent compound, an enzymatic substrate, a radioactive compound or a luminescent compound, or a second antibody which recognizes the first antibody may be conjugated to a detectable substrate).

In another specific embodiment of the present invention, the kit is a diagnostic kit for use in screening serum containing antibodies specific against proliferative and/or cancerous polynucleotides and polypeptides. Such a kit may include a control antibody that does not react with the polypeptide of interest. Such a kit may include a substantially isolated polypeptide antigen comprising an epitope which is specifically immunoreactive with at least one anti-polypeptide antigen antibody. Further, such a kit includes means for detecting the binding of said antibody to the antigen (e.g., the antibody may be conjugated to a fluorescent compound such as fluorescein or rhodamine which can be detected by flow cytometry). In specific embodiments, the kit may include a recombinantly produced or chemically synthesized polypeptide antigen. The polypeptide antigen of the kit may also be attached to a solid support.

In a more specific embodiment the detecting means of the above-described kit includes a solid support to which said polypeptide antigen is attached. Such a kit may also include a non-attached reporter-labeled anti-human antibody. In this embodiment, binding of the antibody to the polypeptide antigen can be detected by binding of the said reporter-labeled antibody.

In an additional embodiment, the invention includes a diagnostic kit for use in screening serum containing antigens of the polypeptide of the invention. The diagnostic kit includes a substantially isolated antibody specifically immunoreactive with polypeptide or polynucleotide antigens, and means for detecting the binding of the polynucleotide or polypeptide antigen to the antibody. In one embodiment, the antibody is attached to a solid support. In a specific embodiment, the antibody may be a monoclonal antibody. The detecting means of the kit may include a second, labeled monoclonal antibody. Alternatively, or in addition, the detecting means may include a labeled, competing antigen.

In one diagnostic configuration, test serum is reacted with a solid phase reagent having a surface-bound antigen obtained by the methods of the present invention. After binding with specific antigen antibody to the reagent and removing unbound serum components by washing, the reagent is reacted with reporter-labeled  
5 anti-human antibody to bind reporter to the reagent in proportion to the amount of bound anti-antigen antibody on the solid support. The reagent is again washed to remove unbound labeled antibody, and the amount of reporter associated with the reagent is determined. Typically, the reporter is an enzyme which is detected by incubating the solid phase in the presence of a suitable fluorometric, luminescent or  
10 colorimetric substrate (Sigma, St. Louis, MO).

The solid surface reagent in the above assay is prepared by known techniques for attaching protein material to solid support material, such as polymeric beads, dip sticks, 96-well plate or filter material. These attachment methods generally include non-specific adsorption of the protein to the support or covalent attachment of the  
15 protein, typically through a free amine group, to a chemically reactive group on the solid support, such as an activated carboxyl, hydroxyl, or aldehyde group. Alternatively, streptavidin coated plates can be used in conjunction with biotinylated antigen(s).

Thus, the invention provides an assay system or kit for carrying out this  
20 diagnostic method. The kit generally includes a support with surface-bound recombinant antigens, and a reporter-labeled anti-human antibody for detecting surface-bound anti-antigen antibody.

#### Fusion Proteins

Any polypeptide of the present invention can be used to generate fusion  
25 proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because certain proteins target cellular locations based on trafficking signals, the polypeptides of the present  
30 invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous

functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. Similarly, peptide cleavage sites can be introduced in-between such peptide moieties, which could additionally be subjected to protease activity to remove said peptide(s) from the protein of the present invention. The addition of peptide moieties, including peptide cleavage sites, to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgA, IgE, IgG, IgM) or portions thereof (CH1, CH2, CH3, and any combination thereof, including both entire domains and portions thereof), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of the constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified,

would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995)). . . . .

Moreover, the polypeptides of the present invention can be fused to marker sequences (also referred to as "tags"). Due to the availability of antibodies specific to such "tags", purification of the fused polypeptide of the invention, and/or its identification is significantly facilitated since antibodies specific to the polypeptides of the invention are not required. Such purification may be in the form of an affinity purification whereby an anti-tag antibody or another type of affinity matrix (e.g., anti-tag antibody attached to the matrix of a flow-thru column) that binds to the epitope tag is present. In preferred embodiments, the marker amino acid sequence is a hexahistidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984)).

The skilled artisan would acknowledge the existence of other "tags" which could be readily substituted for the tags referred to supra for purification and/or identification of polypeptides of the present invention (Jones C., et al., J Chromatogr A. 707(1):3-22 (1995)). For example, the c-myc tag and the 8F9, 3C7, 6E10, G4m B7 and 9E10 antibodies thereto (Evan et al., Molecular and Cellular Biology 5:3610-3616 (1985)); the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky et al., Protein Engineering, 3(6):547-553 (1990), the Flag-peptide – i.e., the octapeptide sequence DYKDDDDK (SEQ ID NO:33), (Hopp et al., Biotech. 6:1204-1210 (1988); the KT3 epitope peptide (Martin et al., Science, 255:192-194 (1992)); a-tubulin epitope peptide (Skinner et al., J. Biol. Chem., 266:15136-15166, (1991)); the T7 gene 10 protein peptide tag (Lutz-Freyermuth et al., Proc. Natl. Sci. USA,

87:6363-6397 (1990)), the FITC epitope (Zymed, Inc.), the GFP epitope (Zymed, Inc.), and the Rhodamine epitope (Zymed, Inc.).

5 The present invention also encompasses the attachment of up to nine codons encoding a repeating series of up to nine arginine amino acids to the coding region of a polynucleotide of the present invention. The invention also encompasses chemically derivitizing a polypeptide of the present invention with a repeating series of up to nine arginine amino acids. Such a tag, when attached to a polypeptide, has recently been shown to serve as a universal pass, allowing compounds access to the interior of cells without additional derivitization or manipulation (Wender, P., et al., unpublished data).

10 Protein fusions involving polypeptides of the present invention, including fragments and/or variants thereof, can be used for the following, non-limiting examples, subcellular localization of proteins, determination of protein-protein interactions via immunoprecipitation, purification of proteins via affinity chromatography, functional and/or structural characterization of protein. The present invention also encompasses the application of hapten specific antibodies for any of the uses referenced above for epitope fusion proteins. For example, the polypeptides of the present invention could be chemically derivatized to attach hapten molecules (e.g., DNP, (Zymed, Inc.)). Due to the availability of monoclonal antibodies specific to such haptens, the protein could be readily purified using immunoprecipitation, for example.

15 Polypeptides of the present invention, including fragments and/or variants thereof, in addition to, antibodies directed against such polypeptides, fragments, and/or variants, may be fused to any of a number of known, and yet to be determined, toxins, such as ricin, saporin (Mashiba H, et al., Ann. N. Y. Acad. Sci. 1999;886:233-5), or HC toxin (Tonukari NJ, et al., Plant Cell. 2000 Feb;12(2):237-248), for example. Such fusions could be used to deliver the toxins to desired tissues for which a ligand or a protein capable of binding to the polypeptides of the invention exists.

20 The invention encompasses the fusion of antibodies directed against polypeptides of the present invention, including variants and fragments thereof, to said toxins for delivering the toxin to specific locations in a cell, to specific tissues, and/or to specific species. Such bifunctional antibodies are known in the art, though a

review describing additional advantageous fusions, including citations for methods of production, can be found in P.J. Hudson, Curr. Opp. In. Imm. 11:548-557, (1999); this publication, in addition to the references cited therein, are hereby incorporated by reference in their entirety herein. In this context, the term "toxin" may be expanded to  
5 include any heterologous protein, a small molecule, radionucleotides, cytotoxic drugs, liposomes, adhesion molecules, glycoproteins, ligands, cell or tissue-specific ligands, enzymes, of bioactive agents, biological response modifiers, anti-fungal agents, hormones, steroids, vitamins, peptides, peptide analogs, anti-allergenic agents, anti-tubercular agents, anti-viral agents, antibiotics, anti-protozoan agents, chelates,  
10 radioactive particles, radioactive ions, X-ray contrast agents, monoclonal antibodies, polyclonal antibodies and genetic material. In view of the present disclosure, one skilled in the art could determine whether any particular "toxin" could be used in the compounds of the present invention. Examples of suitable "toxins" listed above are exemplary only and are not intended to limit the "toxins" that may be used in the  
15 present invention.

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

#### Vectors, Host Cells, and Protein Production

The present invention also relates to vectors containing the polynucleotide of  
20 the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

25 The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

30 The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to

name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells (e.g., *Saccharomyces cerevisiae* or *Pichia pastoris* (ATCC Accession No. 201178)); insect cells such as *Drosophila* S2 and *Spodoptera Sf9* cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Preferred expression vectors for use in yeast systems include, but are not limited to pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalph, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, pPIC9K, and PAO815 (all available from Invitrogen, Carlsbad, CA). Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., *Basic Methods In Molecular Biology* (1986). It is specifically contemplated that the

polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

In one embodiment, the yeast *Pichia pastoris* is used to express the polypeptide of the present invention in a eukaryotic system. *Pichia pastoris* is a methylotrophic yeast which can metabolize methanol as its sole carbon source. A main step in the methanol metabolism pathway is the oxidation of methanol to formaldehyde using O<sub>2</sub>. This reaction is catalyzed by the enzyme alcohol oxidase. In order to metabolize methanol as its sole carbon source, *Pichia pastoris* must generate high levels of alcohol oxidase due, in part, to the relatively low affinity of alcohol oxidase for O<sub>2</sub>. Consequently, in a growth medium depending on methanol as a main

carbon source, the promoter region of one of the two alcohol oxidase genes (AOX1) is highly active. In the presence of methanol, alcohol oxidase produced from the AOX1 gene comprises up to approximately 30% of the total soluble protein in *Pichia pastoris*. See, Ellis, S.B., et al., *Mol. Cell. Biol.* 5:1111-21 (1985); Koutz, P.J., et al.,  
5 *Yeast* 5:167-77 (1989); Tschopp, J.F., et al., *Nucl. Acids Res.* 15:3859-76 (1987). Thus, a heterologous coding sequence, such as, for example, a polynucleotide of the present invention, under the transcriptional regulation of all or part of the AOX1 regulatory sequence is expressed at exceptionally high levels in *Pichia* yeast grown in the presence of methanol.

10 In one example, the plasmid vector pPIC9K is used to express DNA encoding a polypeptide of the invention, as set forth herein, in a *Pichea* yeast system essentially as described in "Pichia Protocols: Methods in Molecular Biology," D.R. Higgins and J. Cregg, eds. The Humana Press, Totowa, NJ, 1998. This expression vector allows expression and secretion of a protein of the invention by virtue of the strong AOX1  
15 promoter linked to the *Pichia pastoris* alkaline phosphatase (PHO) secretory signal peptide (i.e., leader) located upstream of a multiple cloning site.

Many other yeast vectors could be used in place of pPIC9K, such as, pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalpha, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, and PAO815, as one skilled in the art would readily  
20 appreciate, as long as the proposed expression construct provides appropriately located signals for transcription, translation, secretion (if desired), and the like, including an in-frame AUG, as required.

In another embodiment, high-level expression of a heterologous coding sequence, such as, for example, a polynucleotide of the present invention, may be  
25 achieved by cloning the heterologous polynucleotide of the invention into an expression vector such as, for example, pGAPZ or pGAPZalpha, and growing the yeast culture in the absence of methanol.

In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and  
30 immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide

sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination, resulting in the formation of a new transcription unit (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; U.S. Patent No. 5,733,761, issued March 31, 1998; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

In addition, polypeptides of the invention can be chemically synthesized using techniques known in the art (e.g., see Creighton, 1983, Proteins: Structures and Molecular Principles, W.H. Freeman & Co., N.Y., and Hunkapiller et al., Nature, 310:105-111 (1984)). For example, a polypeptide corresponding to a fragment of a polypeptide sequence of the invention can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the polypeptide sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid,  $\alpha$ -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid,  $\gamma$ -Abu,  $\epsilon$ -Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine,  $\beta$ -alanine, fluoro-amino acids, designer amino acids such as  $\beta$ -methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

The invention encompasses polypeptides which are differentially modified during or after translation, e.g., by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not

limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH<sub>4</sub>; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

Additional post-translational modifications encompassed by the invention include, for example, e.g., N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell expression. The polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein, the addition of epitope tagged peptide fragments (e.g., FLAG, HA, GST, thioredoxin, maltose binding protein, etc.), attachment of affinity tags such as biotin and/or streptavidin, the covalent attachment of chemical moieties to the amino acid backbone, N- or C-terminal processing of the polypeptides ends (e.g., proteolytic processing), deletion of the N-terminal methionine residue, etc.

Also provided by the invention are chemically modified derivatives of the polypeptides of the invention which may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Patent NO: 4,179,337). The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

The invention further encompasses chemical derivitization of the polypeptides of the present invention, preferably where the chemical is a hydrophilic polymer residue. Exemplary hydrophilic polymers, including derivatives, may be those that include polymers in which the repeating units contain one or more hydroxy groups (polyhydroxy polymers), including, for example, poly(vinyl alcohol); polymers in which the repeating units contain one or more amino groups (polyamine polymers), including, for example, peptides, polypeptides, proteins and lipoproteins, such as albumin and natural lipoproteins; polymers in which the repeating units contain one or

more carboxy groups (polycarboxy polymers), including, for example, carboxymethylcellulose, alginic acid and salts thereof, such as sodium and calcium alginate, glycosaminoglycans and salts thereof, including salts of hyaluronic acid, phosphorylated and sulfonated derivatives of carbohydrates, genetic material, such as interleukin-2 and interferon, and phosphorothioate oligomers; and polymers in which the repeating units contain one or more saccharide moieties. (polysaccharide polymers), including, for example, carbohydrates.

The molecular weight of the hydrophilic polymers may vary, and is generally about 50 to about 5,000,000, with polymers having a molecular weight of about 100 to about 50,000 being preferred. The polymers may be branched or unbranched. More preferred polymers have a molecular weight of about 150 to about 10,000, with molecular weights of 200 to about 8,000 being even more preferred.

For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

Additional preferred polymers which may be used to derivatize polypeptides of the invention, include, for example, poly(ethylene glycol) (PEG), poly(vinylpyrrolidone), polyoxomers, polysorbate and poly(vinyl alcohol), with PEG polymers being particularly preferred. Preferred among the PEG polymers are PEG polymers having a molecular weight of from about 100 to about 10,000. More preferably, the PEG polymers have a molecular weight of from about 200 to about 8,000, with PEG 2,000, PEG 5,000 and PEG 8,000, which have molecular weights of 2,000, 5,000 and 8,000, respectively, being even more preferred. Other suitable hydrophilic polymers, in addition to those exemplified above, will be readily apparent to one skilled in the art based on the present disclosure. Generally, the polymers used may include polymers that can be attached to the polypeptides of the invention via alkylation or acylation reactions.

The polyethylene glycol molecules (or other chemical moieties) should be attached to the protein with consideration of effects on functional or antigenic domains of the protein. There are a number of attachment methods available to those skilled in the art, e.g., EP 0 401 384, herein incorporated by reference (coupling PEG to G-CSF), see also Malik et al., *Exp. Hematol.* 20:1028-1035 (1992) (reporting pegylation of GM-CSF using tresyl chloride). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as, a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. Preferred for therapeutic purposes is attachment at an amino group, such as attachment at the N-terminus or lysine group.

One may specifically desire proteins chemically modified at the N-terminus. Using polyethylene glycol as an illustration of the present composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, etc.), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus modification may be accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminus) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

As with the various polymers exemplified above, it is contemplated that the polymeric residues may contain functional groups in addition, for example, to those

typically involved in linking the polymeric residues to the polypeptides of the present invention. Such functionalities include, for example, carboxyl, amine, hydroxy and thiol groups. These functional groups on the polymeric residues can be further reacted, if desired, with materials that are generally reactive with such functional groups and which can assist in targeting specific tissues in the body including, for example, diseased tissue. Exemplary materials which can be reacted with the additional functional groups include, for example, proteins, including antibodies, carbohydrates, peptides, glycopeptides, glycolipids, lectins, and nucleosides.

In addition to residues of hydrophilic polymers, the chemical used to derivatize the polypeptides of the present invention can be a saccharide residue. Exemplary saccharides which can be derived include, for example, monosaccharides or sugar alcohols, such as erythrose, threose, ribose, arabinose, xylose, lyxose, fructose, sorbitol, mannitol and sedoheptulose, with preferred monosaccharides being fructose, mannose, xylose, arabinose, mannitol and sorbitol; and disaccharides, such as lactose, sucrose, maltose and cellobiose. Other saccharides include, for example, inositol and ganglioside head groups. Other suitable saccharides, in addition to those exemplified above, will be readily apparent to one skilled in the art based on the present disclosure. Generally, saccharides which may be used for derivitization include saccharides that can be attached to the polypeptides of the invention via alkylation or acylation reactions.

Moreover, the invention also encompasses derivitization of the polypeptides of the present invention, for example, with lipids (including cationic, anionic, polymerized, charged, synthetic, saturated, unsaturated, and any combination of the above, etc.). stabilizing agents.

The invention encompasses derivitization of the polypeptides of the present invention, for example, with compounds that may serve a stabilizing function (e.g., to increase the polypeptides half-life in solution, to make the polypeptides more water soluble, to increase the polypeptides hydrophilic or hydrophobic character, etc.). Polymers useful as stabilizing materials may be of natural, semi-synthetic (modified natural) or synthetic origin. Exemplary natural polymers include naturally occurring polysaccharides, such as, for example, arabinans, fructans, fucans, galactans, galacturonans, glucans, mannans, xylans (such as, for example, inulin), levan,

fucoïdan, carrageenan, galatocarolose, pectic acid, pectins, including amylose, pullulan, glycogen, amylopectin, cellulose, dextran, dextrin, dextrose, glucose, polyglucose, polydextrose, pustulan, chitin, agarose, keratin, chondroitin, dermatan, hyaluronic acid, alginic acid, xanthin gum, starch and various other natural

5 homopolymer or heteropolymers, such as those containing one or more of the following aldoses, ketoses, acids or amines: erythrose, threose, ribose, arabinose, xylose, lyxose, allose, altrose, glucose, dextrose, mannose, gulose, idose, galactose, talose, erythrulose, ribulose, xylulose, psicose, fructose, sorbose, tagatose, mannitol, sorbitol, lactose, sucrose, trehalose, maltose, cellobiose, glycine, serine, threonine,

10 cysteine, tyrosine, asparagine, glutamine, aspartic acid, glutamic acid, lysine, arginine, histidine, glucuronic acid, gluconic acid, glucaric acid, galacturonic acid, mannuronic acid, glucosamine, galactosamine, and neuraminic acid, and naturally occurring derivatives thereof Accordingly, suitable polymers include, for example, proteins, such as albumin, polyalginates, and polylactide-coglycolide polymers.

15 Exemplary semi-synthetic polymers include carboxymethylcellulose, hydroxymethylcellulose, hydroxypropylmethylcellulose, methylcellulose, and methoxycellulose. Exemplary synthetic polymers include polyphosphazenes, hydroxyapatites, fluoroapatite polymers, polyethylenes (such as, for example, polyethylene glycol (including for example, the class of compounds referred to as

20 Pluronic<sup>®</sup> RTM., commercially available from BASF, Parsippany, N.J.), polyoxyethylene, and polyethylene terephthalate), polypropylenes (such as, for example, polypropylene glycol), polyurethanes (such as, for example, polyvinyl alcohol (PVA), polyvinyl chloride and polyvinylpyrrolidone), polyamides including nylon, polystyrene, polylactic acids, fluorinated hydrocarbon polymers, fluorinated

25 carbon polymers (such as, for example, polytetrafluoroethylene), acrylate, methacrylate, and polymethylmethacrylate, and derivatives thereof. Methods for the preparation of derivatized polypeptides of the invention which employ polymers as stabilizing compounds will be readily apparent to one skilled in the art, in view of the present disclosure, when coupled with information known in the art, such as that

30 described and referred to in Unger, U.S. Pat. No. 5,205,290, the disclosure of which is hereby incorporated by reference herein in its entirety.

Moreover, the invention encompasses additional modifications of the polypeptides of the present invention. Such additional modifications are known in the art, and are specifically provided, in addition to methods of derivitization, etc., in US Patent No. 6,028,066, which is hereby incorporated in its entirety herein.

5       The polypeptides of the invention may be in monomers or multimers (i.e., dimers, trimers, tetramers and higher multimers). Accordingly, the present invention relates to monomers and multimers of the polypeptides of the invention, their preparation, and compositions (preferably, Therapeutics) containing them. In specific embodiments, the polypeptides of the invention are monomers, dimers, trimers or  
10       tetramers. In additional embodiments, the multimers of the invention are at least dimers, at least trimers, or at least tetramers.

      Multimers encompassed by the invention may be homomers or heteromers. As used herein, the term homomer, refers to a multimer containing only polypeptides corresponding to the amino acid sequence of SEQ ID NO:2 or encoded by the cDNA  
15       contained in a deposited clone (including fragments, variants, splice variants, and fusion proteins, corresponding to these polypeptides as described herein). These homomers may contain polypeptides having identical or different amino acid sequences. In a specific embodiment, a homomer of the invention is a multimer containing only polypeptides having an identical amino acid sequence. In another  
20       specific embodiment, a homomer of the invention is a multimer containing polypeptides having different amino acid sequences. In specific embodiments, the multimer of the invention is a homodimer (e.g., containing polypeptides having identical or different amino acid sequences) or a homotrimer (e.g., containing polypeptides having identical and/or different amino acid sequences). In additional  
25       embodiments, the homomeric multimer of the invention is at least a homodimer, at least a homotrimer, or at least a homotetramer.

      As used herein, the term heteromer refers to a multimer containing one or more heterologous polypeptides (i.e., polypeptides of different proteins) in addition to the polypeptides of the invention. In a specific embodiment, the multimer of the  
30       invention is a heterodimer, a heterotrimer, or a heterotetramer. In additional embodiments, the heteromeric multimer of the invention is at least a heterodimer, at least a heterotrimer, or at least a heterotetramer.

Multimers of the invention may be the result of hydrophobic, hydrophilic, ionic and/or covalent associations and/or may be indirectly linked, by for example, liposome formation. Thus, in one embodiment, multimers of the invention, such as, for example, homodimers or homotrimers, are formed when polypeptides of the invention contact one another in solution. In another embodiment, heteromultimers of the invention, such as, for example, heterotrimers or heterotetramers, are formed when polypeptides of the invention contact antibodies to the polypeptides of the invention (including antibodies to the heterologous polypeptide sequence in a fusion protein of the invention) in solution. In other embodiments, multimers of the invention are formed by covalent associations with and/or between the polypeptides of the invention. Such covalent associations may involve one or more amino acid residues contained in the polypeptide sequence (e.g., that recited in the sequence listing, or contained in the polypeptide encoded by a deposited clone). In one instance, the covalent associations are cross-linking between cysteine residues located within the polypeptide sequences which interact in the native (i.e., naturally occurring) polypeptide. In another instance, the covalent associations are the consequence of chemical or recombinant manipulation. Alternatively, such covalent associations may involve one or more amino acid residues contained in the heterologous polypeptide sequence in a fusion protein of the invention.

In one example, covalent associations are between the heterologous sequence contained in a fusion protein of the invention (see, e.g., US Patent Number 5,478,925). In a specific example, the covalent associations are between the heterologous sequence contained in an Fc fusion protein of the invention (as described herein). In another specific example, covalent associations of fusion proteins of the invention are between heterologous polypeptide sequence from another protein that is capable of forming covalently associated multimers, such as for example, osteoprotegerin (see, e.g., International Publication NO: WO 98/49305, the contents of which are herein incorporated by reference in its entirety). In another embodiment, two or more polypeptides of the invention are joined through peptide linkers. Examples include those peptide linkers described in U.S. Pat. No. 5,073,627 (hereby incorporated by reference). Proteins comprising multiple polypeptides of the

invention separated by peptide linkers may be produced using conventional recombinant DNA technology.

Another method for preparing multimer polypeptides of the invention involves use of polypeptides of the invention fused to a leucine zipper or isoleucine zipper polypeptide sequence. Leucine zipper and isoleucine zipper domains are polypeptides that promote multimerization of the proteins in which they are found. Leucine zippers were originally identified in several DNA-binding proteins (Landschulz et al., Science 240:1759, (1988)), and have since been found in a variety of different proteins. Among the known leucine zippers are naturally occurring peptides and derivatives thereof that dimerize or trimerize. Examples of leucine zipper domains suitable for producing soluble multimeric proteins of the invention are those described in PCT application WO 94/10308, hereby incorporated by reference. Recombinant fusion proteins comprising a polypeptide of the invention fused to a polypeptide sequence that dimerizes or trimerizes in solution are expressed in suitable host cells, and the resulting soluble multimeric fusion protein is recovered from the culture supernatant using techniques known in the art.

Trimeric polypeptides of the invention may offer the advantage of enhanced biological activity. Preferred leucine zipper moieties and isoleucine moieties are those that preferentially form trimers. One example is a leucine zipper derived from lung surfactant protein D (SPD), as described in Hoppe et al. (FEBS Letters 344:191, (1994)) and in U.S. patent application Ser. No. 08/446,922, hereby incorporated by reference. Other peptides derived from naturally occurring trimeric proteins may be employed in preparing trimeric polypeptides of the invention.

In another example, proteins of the invention are associated by interactions between Flag® polypeptide sequence contained in fusion proteins of the invention containing Flag® polypeptide sequence. In a further embodiment, associations proteins of the invention are associated by interactions between heterologous polypeptide sequence contained in Flag® fusion proteins of the invention and anti-Flag® antibody.

The multimers of the invention may be generated using chemical techniques known in the art. For example, polypeptides desired to be contained in the multimers of the invention may be chemically cross-linked using linker molecules and linker

molecule length optimization techniques known in the art (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Additionally, multimers of the invention may be generated using techniques known in the art to form one or more inter-molecule cross-links between the cysteine residues located within the sequence of the polypeptides desired to be contained in the multimer (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Further, polypeptides of the invention may be routinely modified by the addition of cysteine or biotin to the C terminus or N-terminus of the polypeptide and techniques known in the art may be applied to generate multimers containing one or more of these modified polypeptides (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Additionally, techniques known in the art may be applied to generate liposomes containing the polypeptide components desired to be contained in the multimer of the invention (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety).

Alternatively, multimers of the invention may be generated using genetic engineering techniques known in the art. In one embodiment, polypeptides contained in multimers of the invention are produced recombinantly using fusion protein technology described herein or otherwise known in the art (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). In a specific embodiment, polynucleotides coding for a homodimer of the invention are generated by ligating a polynucleotide sequence encoding a polypeptide of the invention to a sequence encoding a linker polypeptide and then further to a synthetic polynucleotide encoding the translated product of the polypeptide in the reverse orientation from the original C-terminus to the N-terminus (lacking the leader sequence) (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). In another embodiment, recombinant techniques described herein or otherwise known in the art are applied to generate recombinant polypeptides of the invention which contain a transmembrane domain (or hydrophobic or signal peptide) and which can be incorporated by membrane reconstitution techniques into liposomes (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety).

In addition, the polynucleotide insert of the present invention could be operatively linked to "artificial" or chimeric promoters and transcription factors. Specifically, the artificial promoter could comprise, or alternatively consist, of any combination of cis-acting DNA sequence elements that are recognized by trans-acting transcription factors. Preferably, the cis acting DNA sequence elements and trans-acting transcription factors are operable in mammals. Further, the trans-acting transcription factors of such "artificial" promoters could also be "artificial" or chimeric in design themselves and could act as activators or repressors to said "artificial" promoter.

#### 10 Uses of the Polynucleotides

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:1. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:1 will yield an amplified fragment.

Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see  
5 Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides  
10 correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage  
15 analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. Disease mapping data are known in the art. Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

20 Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected organisms can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed  
25 in some or all affected organisms, but not in normal organisms, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal organisms is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

30 Furthermore, increased or decreased expression of the gene in affected organisms as compared to unaffected organisms can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression,

chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

Thus, the invention also provides a diagnostic method useful during diagnosis of a disorder, involving measuring the expression level of polynucleotides of the present invention in cells or body fluid from an organism and comparing the  
5 measured gene expression level with a standard level of polynucleotide expression level, whereby an increase or decrease in the gene expression level compared to the standard is indicative of a disorder.

By "measuring the expression level of a polynucleotide of the present  
10 invention" is intended qualitatively or quantitatively measuring or estimating the level of the polypeptide of the present invention or the level of the mRNA encoding the polypeptide in a first biological sample either directly (e.g., by determining or estimating absolute protein level or mRNA level) or relatively (e.g., by comparing to the polypeptide level or mRNA level in a second biological sample). Preferably, the  
15 polypeptide level or mRNA level in the first biological sample is measured or estimated and compared to a standard polypeptide level or mRNA level, the standard being taken from a second biological sample obtained from an individual not having the disorder or being determined by averaging levels from a population of organisms not having a disorder. As will be appreciated in the art, once a standard polypeptide  
20 level or mRNA level is known, it can be used repeatedly as a standard for comparison.

By "biological sample" is intended any biological sample obtained from an organism, body fluids, cell line, tissue culture, or other source which contains the polypeptide of the present invention or mRNA. As indicated, biological samples  
25 include body fluids (such as the following non-limiting examples, sputum, amniotic fluid, urine, saliva, breast milk, secretions, interstitial fluid, blood, serum, spinal fluid, etc.) which contain the polypeptide of the present invention, and other tissue sources found to express the polypeptide of the present invention. Methods for obtaining tissue biopsies and body fluids from organisms are well known in the art. Where the  
30 biological sample is to include mRNA, a tissue biopsy is the preferred source.

The method(s) provided above may Preferably be applied in a diagnostic method and/or kits in which polynucleotides and/or polypeptides are attached to a

solid support. In one exemplary method, the support may be a "gene chip" or a "biological chip" as described in US Patents 5,837,832, 5,874,219, and 5,856,174. Further, such a gene chip with polynucleotides of the present invention attached may be used to identify polymorphisms between the polynucleotide sequences, with  
5 polynucleotides isolated from a test subject. The knowledge of such polymorphisms (i.e. their location, as well as, their existence) would be beneficial in identifying disease loci for many disorders, including proliferative diseases and conditions. Such a method is described in US Patents 5,858,659 and 5,856,104. The US Patents referenced supra are hereby incorporated by reference in their entirety herein.

10 The present invention encompasses polynucleotides of the present invention that are chemically synthesized, or reproduced as peptide nucleic acids (PNA), or according to other methods known in the art. The use of PNAs would serve as the preferred form if the polynucleotides are incorporated onto a solid support, or gene chip. For the purposes of the present invention, a peptide nucleic acid (PNA) is a  
15 polyamide type of DNA analog and the monomeric units for adenine, guanine, thymine and cytosine are available commercially (Perceptive Biosystems). Certain components of DNA, such as phosphorus, phosphorus oxides, or deoxyribose derivatives, are not present in PNAs. As disclosed by P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, Science 254, 1497 (1991); and M. Egholm, O. Buchardt,  
20 L.Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, and P. E. Nielsen, Nature 365, 666 (1993), PNAs bind specifically and tightly to complementary DNA strands and are not degraded by nucleases. In fact, PNA binds more strongly to DNA than DNA itself does. This is probably because there is no electrostatic repulsion between the two strands, and also the polyamide  
25 backbone is more flexible. Because of this, PNA/DNA duplexes bind under a wider range of stringency conditions than DNA/DNA duplexes, making it easier to perform multiplex hybridization. Smaller probes can be used than with DNA due to the stronger binding characteristics of PNA:DNA hybrids. In addition, it is more likely that single base mismatches can be determined with PNA/DNA hybridization because  
30 a single mismatch in a PNA/DNA 15-mer lowers the melting point ( $T_{sub.m}$ ) by 8°-20° C, vs. 4°-16° C for the DNA/DNA 15-mer duplex. Also, the absence of charge

groups in PNA means that hybridization can be done at low ionic strengths and reduce possible interference by salt during the analysis.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Antisense techniques are discussed, for example, in Okano, J. *Neurochem.* 56: 560 (1991); "Oligodeoxynucleotides as Antisense. Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Triple helix formation is discussed in, for instance Lee et al., *Nucleic Acids Research* 6: 3073 (1979); Cooney et al., *Science* 241: 456 (1988); and Dervan et al., *Science* 251: 1360 (1991). Both methods rely on binding of the polynucleotide to a complementary DNA or RNA. For these techniques, preferred polynucleotides are usually oligonucleotides 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., *Nucl. Acids Res.* 6:3073 (1979); Cooney et al., *Science* 241:456 (1988); and Dervan et al., *Science* 251:1360 (1991) ) or to the mRNA itself (antisense - Okano, J. *Neurochem.* 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat or prevent disease.

The present invention encompasses the addition of a nuclear localization signal, operably linked to the 5' end, 3' end, or any location therein, to any of the oligonucleotides, antisense oligonucleotides, triple helix oligonucleotides, ribozymes, PNA oligonucleotides, and/or polynucleotides, of the present invention. See, for example, G. Cutrona, et al., *Nat. Biotech.*, 18:300-303, (2000); which is hereby incorporated herein by reference.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome,

thereby producing a new trait in the host cell. In one example, polynucleotide sequences of the present invention may be used to construct chimeric RNA/DNA oligonucleotides corresponding to said sequences, specifically designed to induce host cell mismatch repair mechanisms in an organism upon systemic injection, for example  
5 (Bartlett, R.J., et al., Nat. Biotech, 18:615-622 (2000), which is hereby incorporated by reference herein in its entirety). Such RNA/DNA oligonucleotides could be designed to correct genetic defects in certain host strains, and/or to introduce desired phenotypes in the host (e.g., introduction of a specific polymorphism within an endogenous gene corresponding to a polynucleotide of the present invention that may  
10 ameliorate and/or prevent a disease symptom and/or disorder, etc.). Alternatively, the polynucleotide sequence of the present invention may be used to construct duplex oligonucleotides corresponding to said sequence, specifically designed to correct genetic defects in certain host strains, and/or to introduce desired phenotypes into the host (e.g., introduction of a specific polymorphism within an endogenous gene  
15 corresponding to a polynucleotide of the present invention that may ameliorate and/or prevent a disease symptom and/or disorder, etc.). Such methods of using duplex oligonucleotides are known in the art and are encompassed by the present invention (see EP1007712, which is hereby incorporated by reference herein in its entirety).

The polynucleotides are also useful for identifying organisms from minute  
20 biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of  
25 "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions  
30 of an organisms genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, organisms can be identified because each organism will have a unique set

of DNA sequences. Once an unique ID database is established for an organism, positive identification of that organism, living or dead, can be made from extremely small tissue samples. Similarly, polynucleotides of the present invention can be used as polymorphic markers, in addition to, the identification of transformed or non-transformed cells and/or tissues.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination. Moreover, as mentioned above, such reagents can be used to screen and/or identify transformed and non-transformed cells and/or tissues.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

#### Uses of the Polypeptides

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and

technetium ( $^{99m}\text{Tc}$ ), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example,  $^{131}\text{I}$ ,  $^{112}\text{In}$ ,  $^{99m}\text{Tc}$ ), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of  $^{99m}\text{Tc}$ . The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder. With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more

definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Moreover, polypeptides of the present invention can be used to treat, prevent, and/or diagnose disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B, SOD, catalase, DNA repair proteins), to inhibit the activity of a polypeptide (e.g., an oncogene or tumor suppressor), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth inhibition, enhancement of the immune response to proliferative cells or tissues).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat, prevent, and/or diagnose disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

#### Gene Therapy Methods

Another aspect of the present invention is to gene therapy methods for treating or preventing disorders, diseases and conditions. The gene therapy methods relate to the introduction of nucleic acid (DNA, RNA and antisense DNA or RNA) sequences into an animal to achieve expression of a polypeptide of the present invention. This

method requires a polynucleotide which codes for a polypeptide of the invention that operatively linked to a promoter and any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques are known in the art, see, for example, WO90/11092, which is herein  
5 incorporated by reference.

Thus, for example, cells from a patient may be engineered with a polynucleotide (DNA or RNA) comprising a promoter operably linked to a polynucleotide of the invention ex vivo, with the engineered cells then being provided to a patient to be treated with the polypeptide. Such methods are well-known in the  
10 art. For example, see Belldegrün et al., J. Natl. Cancer Inst., 85:207-216 (1993); Ferrantini et al., Cancer Research, 53:107-1112 (1993); Ferrantini et al., J. Immunology 153: 4604-4615 (1994); Kaido, T., et al., Int. J. Cancer 60: 221-229 (1995); Ogura et al., Cancer Research 50: 5102-5106 (1990); Santodonato, et al., Human Gene Therapy 7:1-10 (1996); Santodonato, et al., Gene Therapy 4:1246-1255  
15 (1997); and Zhang, et al., Cancer Gene Therapy 3: 31-38 (1996)), which are herein incorporated by reference. In one embodiment, the cells which are engineered are arterial cells. The arterial cells may be reintroduced into the patient through direct injection to the artery, the tissues surrounding the artery, or through catheter injection.

As discussed in more detail below, the polynucleotide constructs can be  
20 delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, and the like). The polynucleotide constructs may be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

In one embodiment, the polynucleotide of the invention is delivered as a naked  
25 polynucleotide. The term "naked" polynucleotide, DNA or RNA refers to sequences that are free from any delivery vehicle that acts to assist, promote or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the invention can also be delivered in liposome formulations and lipofectin formulations  
30 and the like can be prepared by methods well known to those skilled in the art. Such methods are described, for example, in U.S. Patent Nos. 5,593,972, 5,589,466, and 5,580,859, which are herein incorporated by reference.

The polynucleotide vector constructs of the invention used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Appropriate vectors include pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; pSVK3, 5 pBPV, pMSG and pSVL available from Pharmacia; and pEF1/V5, pcDNA3.1, and pRc/CMV2 available from Invitrogen. Other suitable vectors will be readily apparent to the skilled artisan.

Any strong promoter known to those skilled in the art can be used for driving the expression of polynucleotide sequence of the invention. Suitable promoters 10 include adenoviral promoters, such as the adenoviral major late promoter; or heterologous promoters, such as the cytomegalovirus (CMV) promoter; the respiratory syncytial virus (RSV) promoter; inducible promoters, such as the MMT promoter, the metallothionein promoter; heat shock promoters; the albumin promoter; the ApoAI promoter; human globin promoters; viral thymidine kinase promoters, 15 such as the Herpes Simplex thymidine kinase promoter; retroviral LTRs; the b-actin promoter; and human growth hormone promoters. The promoter also may be the native promoter for the polynucleotides of the invention.

Unlike other gene therapy techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the 20 polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct of the invention can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, 25 lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular, fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen 30 fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the

interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-  
5 differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. In vivo muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked nucleic acid sequence injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 mg/kg body weight to about 50  
10 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the  
15 condition being treated and the route of administration.

The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked DNA  
20 constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The naked polynucleotides are delivered by any method known in the art, including, but not limited to, direct needle injection at the delivery site, intravenous injection, topical administration, catheter infusion, and so-called "gene guns". These  
25 delivery methods are known in the art.

The constructs may also be delivered with delivery vehicles such as viral sequences, viral particles, liposome formulations, lipofectin, precipitating agents, etc. Such methods of delivery are known in the art.

In certain embodiments, the polynucleotide constructs of the invention are  
30 complexed in a liposome preparation. Liposomal preparations for use in the instant invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. However, cationic liposomes are particularly preferred because a

tight charge complex can be formed between the cationic liposome and the polyanionic nucleic acid. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner et al., Proc. Natl. Acad. Sci. USA , 84:7413-7416 (1987), which is herein incorporated by reference); mRNA (Malone et al., Proc. Natl. Acad. Sci. USA , 86:6077-6081 (1989), which is herein incorporated by reference); and purified transcription factors (Debs et al., J. Biol. Chem., 265:10189-10192 (1990), which is herein incorporated by reference), in functional form.

Cationic liposomes are readily available. For example, N[1-2,3-dioleoyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are particularly useful and are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, N.Y. (See, also, Felgner et al., Proc. Natl. Acad. Sci. USA , 84:7413-7416 (1987), which is herein incorporated by reference). Other commercially available liposomes include transfectace (DDAB/DOPE) and DOTAP/DOPE (Boehringer).

Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, e.g. PCT Publication NO: WO 90/11092 (which is herein incorporated by reference) for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes. Preparation of DOTMA liposomes is explained in the literature, see, e.g., Felgner et al., Proc. Natl. Acad. Sci. USA, 84:7413-7417, which is herein incorporated by reference. Similar methods can be used to prepare liposomes from other cationic lipid materials.

Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, Ala.), or can be easily prepared using readily available materials. Such materials include phosphatidyl, choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphosphatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

For example, commercially dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), and dioleoylphosphatidyl ethanolamine (DOPE) can be used in various combinations to make conventional liposomes, with or

without the addition of cholesterol. Thus, for example, DOPG/DOPC vesicles can be prepared by drying 50 mg each of DOPG and DOPC under a stream of nitrogen gas into a sonication vial. The sample is placed under a vacuum pump overnight and is hydrated the following day with deionized water. The sample is then sonicated for 2  
5 hours in a capped vial, using a Heat Systems model 350 sonicator equipped with an inverted cup (bath type) probe at the maximum setting while the bath is circulated at 15EC. Alternatively, negatively charged vesicles can be prepared without sonication to produce multilamellar vesicles or by extrusion through nucleopore membranes to produce unilamellar vesicles of discrete size. Other methods are known and available  
10 to those of skill in the art.

The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs), with SUVs being preferred. The various liposome-nucleic acid complexes are prepared using methods well known in the art. See, e.g., Straubinger et al., *Methods of Immunology*, 101:512-527 (1983),  
15 which is herein incorporated by reference. For example, MLVs containing nucleic acid can be prepared by depositing a thin film of phospholipid on the walls of a glass tube and subsequently hydrating with a solution of the material to be encapsulated. SUVs are prepared by extended sonication of MLVs to produce a homogeneous population of unilamellar liposomes. The material to be entrapped is added to a  
20 suspension of preformed MLVs and then sonicated. When using liposomes containing cationic lipids, the dried lipid film is resuspended in an appropriate solution such as sterile water or an isotonic buffer solution such as 10 mM Tris/NaCl, sonicated, and then the preformed liposomes are mixed directly with the DNA. The liposome and DNA form a very stable complex due to binding of the positively charged liposomes  
25 to the cationic DNA. SUVs find use with small nucleic acid fragments. LUVs are prepared by a number of methods, well known in the art. Commonly used methods include  $\text{Ca}^{2+}$ -EDTA chelation (Papahadjopoulos et al., *Biochim. Biophys. Acta*, 394:483 (1975); Wilson et al., *Cell*, 17:77 (1979)); ether injection (Deamer et al., *Biochim. Biophys. Acta*, 443:629 (1976); Ostro et al., *Biochem. Biophys. Res.*  
30 *Commun.*, 76:836 (1977); Fraley et al., *Proc. Natl. Acad. Sci. USA*, 76:3348 (1979)); detergent dialysis (Enoch et al., *Proc. Natl. Acad. Sci. USA*, 76:145 (1979)); and reverse-phase evaporation (REV) (Fraley et al., *J. Biol. Chem.*, 255:10431 (1980));

Szoka et al., Proc. Natl. Acad. Sci. USA , 75:145 (1978); Schaefer-Ridder et al., Science, 215:166 (1982)), which are herein incorporated by reference.

Generally, the ratio of DNA to liposomes will be from about 10:1 to about 1:10. Preferably, the ration will be from about 5:1 to about 1:5. More preferably, the  
5 ration will be about 3:1 to about 1:3. Still more preferably, the ratio will be about 1:1.

U.S. Patent NO: 5,676,954 (which is herein incorporated by reference) reports on the injection of genetic material, complexed with cationic liposomes carriers, into mice. U.S. Patent Nos. 4,897,355, 4,946,787, 5,049,386, 5,459,127, 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication NO: WO 94/9469  
10 (which are herein incorporated by reference) provide cationic lipids for use in transfecting DNA into cells and mammals. U.S. Patent Nos. 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication NO: WO 94/9469 (which are herein incorporated by reference) provide methods for delivering DNA-cationic lipid complexes to mammals.

15 In certain embodiments, cells are engineered, ex vivo or in vivo, using a retroviral particle containing RNA which comprises a sequence encoding polypeptides of the invention. Retroviruses from which the retroviral plasmid vectors may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, Rous sarcoma Virus, Harvey Sarcoma Virus, avian leukosis  
20 virus, gibbon ape leukemia virus, human immunodeficiency virus, Myeloproliferative Sarcoma Virus, and mammary tumor virus.

The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PE501, PA317, R-2, R-AM, PA12, T19-14X, VT-  
25 19-17-H2, RCRE, RCRIP, GP+E-86, GP+envAm12, and DAN cell lines as described in Miller, Human Gene Therapy , 1:5-14 (1990), which is incorporated herein by reference in its entirety. The vector may transduce the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and CaPO<sub>4</sub> precipitation. In one alternative, the retroviral  
30 plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line generates infectious retroviral vector particles which include polynucleotide encoding polypeptides of the invention. Such retroviral vector particles then may be employed, to transduce eukaryotic cells, either in vitro or in vivo. The transduced eukaryotic cells will express polypeptides of the invention.

5 In certain other embodiments, cells are engineered, ex vivo or in vivo, with polynucleotides of the invention contained in an adenovirus vector. Adenovirus can be manipulated such that it encodes and expresses polypeptides of the invention, and at the same time is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. Adenovirus expression is achieved without integration of the viral DNA  
10 into the host cell chromosome, thereby alleviating concerns about insertional mutagenesis. Furthermore, adenoviruses have been used as live enteric vaccines for many years with an excellent safety profile (Schwartz et al., Am. Rev. Respir. Dis., 109:233-238 (1974)). Finally, adenovirus mediated gene transfer has been demonstrated in a number of instances including transfer of alpha-1-antitrypsin and  
15 CFTR to the lungs of cotton rats (Rosenfeld et al., Science, 252:431-434 (1991); Rosenfeld et al., Cell, 68:143-155 (1992)). Furthermore, extensive studies to attempt to establish adenovirus as a causative agent in human cancer were uniformly negative (Green et al. Proc. Natl. Acad. Sci. USA, 76:6606 (1979)).

Suitable adenoviral vectors useful in the present invention are described, for  
20 example, in Kozarsky and Wilson, Curr. Opin. Genet. Devel., 3:499-503 (1993); Rosenfeld et al., Cell, 68:143-155 (1992); Engelhardt et al., Human Genet. Ther., 4:759-769 (1993); Yang et al., Nature Genet., 7:362-369 (1994); Wilson et al., Nature, 365:691-692 (1993); and U.S. Patent NO: 5,652,224, which are herein incorporated by reference. For example, the adenovirus vector Ad2 is useful and can be grown in  
25 human 293 cells. These cells contain the E1 region of adenovirus and constitutively express Ela and Elb, which complement the defective adenoviruses by providing the products of the genes deleted from the vector. In addition to Ad2, other varieties of adenovirus (e.g., Ad3, Ad5, and Ad7) are also useful in the present invention.

Preferably, the adenoviruses used in the present invention are replication  
30 deficient. Replication deficient adenoviruses require the aid of a helper virus and/or packaging cell line to form infectious particles. The resulting virus is capable of infecting cells and can express a polynucleotide of interest which is operably linked to

a promoter, but cannot replicate in most cells. Replication deficient adenoviruses may be deleted in one or more of all or a portion of the following genes: E1a, E1b, E3, E4, E2a, or L1 through L5.

5 In certain other embodiments, the cells are engineered, ex vivo or in vivo, using an adeno-associated virus (AAV). AAVs are naturally occurring defective viruses that require helper viruses to produce infectious particles (Muzyczka, Curr. Topics in Microbiol. Immunol., 158:97 (1992)). It is also one of the few viruses that may integrate its DNA into non-dividing cells. Vectors containing as little as 300 base pairs of AAV can be packaged and can integrate, but space for exogenous DNA is  
10 limited to about 4.5 kb. Methods for producing and using such AAVs are known in the art. See, for example, U.S. Patent Nos. 5,139,941, 5,173,414, 5,354,678, 5,436,146, 5,474,935, 5,478,745, and 5,589,377.

For example, an appropriate AAV vector for use in the present invention will include all the sequences necessary for DNA replication, encapsidation, and host-cell  
15 integration. The polynucleotide construct containing polynucleotides of the invention is inserted into the AAV vector using standard cloning methods, such as those found in Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press (1989). The recombinant AAV vector is then transfected into packaging cells which are infected with a helper virus, using any standard technique, including  
20 lipofection, electroporation, calcium phosphate precipitation, etc. Appropriate helper viruses include adenoviruses, cytomegaloviruses, vaccinia viruses, or herpes viruses. Once the packaging cells are transfected and infected, they will produce infectious AAV viral particles which contain the polynucleotide construct of the invention. These viral particles are then used to transduce eukaryotic cells, either ex vivo or in  
25 vivo. The transduced cells will contain the polynucleotide construct integrated into its genome, and will express the desired gene product.

Another method of gene therapy involves operably associating heterologous control regions and endogenous polynucleotide sequences (e.g. encoding the polypeptide sequence of interest) via homologous recombination (see, e.g., U.S.  
30 Patent NO: 5,641,670, issued June 24, 1997; International Publication NO: WO 96/29411, published September 26, 1996; International Publication NO: WO 94/12650, published August 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA,

86:8932-8935 (1989); and Zijlstra et al., Nature, 342:435-438 (1989). This method involves the activation of a gene which is present in the target cells, but which is not normally expressed in the cells, or is expressed at a lower level than desired.

Polynucleotide constructs are made, using standard techniques known in the art, which contain the promoter with targeting sequences flanking the promoter. Suitable promoters are described herein. The targeting sequence is sufficiently complementary to an endogenous sequence to permit homologous recombination of the promoter-targeting sequence with the endogenous sequence. The targeting sequence will be sufficiently near the 5' end of the desired endogenous polynucleotide sequence so the promoter will be operably linked to the endogenous sequence upon homologous recombination.

The promoter and the targeting sequences can be amplified using PCR. Preferably, the amplified promoter contains distinct restriction enzyme sites on the 5' and 3' ends. Preferably, the 3' end of the first targeting sequence contains the same restriction enzyme site as the 5' end of the amplified promoter and the 5' end of the second targeting sequence contains the same restriction site as the 3' end of the amplified promoter. The amplified promoter and targeting sequences are digested and ligated together.

The promoter-targeting sequence construct is delivered to the cells, either as naked polynucleotide, or in conjunction with transfection-facilitating agents, such as liposomes, viral sequences, viral particles, whole viruses, lipofection, precipitating agents, etc., described in more detail above. The P promoter-targeting sequence can be delivered by any method, included direct needle injection, intravenous injection, topical administration, catheter infusion, particle accelerators, etc. The methods are described in more detail below.

The promoter-targeting sequence construct is taken up by cells. Homologous recombination between the construct and the endogenous sequence takes place, such that an endogenous sequence is placed under the control of the promoter. The promoter then drives the expression of the endogenous sequence.

The polynucleotides encoding polypeptides of the present invention may be administered along with other polynucleotides encoding angiogenic proteins. Angiogenic proteins include, but are not limited to, acidic and basic fibroblast growth

factors, VEGF-1, VEGF-2 (VEGF-C), VEGF-3 (VEGF-B), epidermal growth factor alpha and beta, platelet-derived endothelial cell growth factor, platelet-derived growth factor, tumor necrosis factor alpha, hepatocyte growth factor, insulin like growth factor, colony stimulating factor, macrophage colony stimulating factor, 5 granulocyte/macrophage colony stimulating factor, and nitric oxide synthase.

Preferably, the polynucleotide encoding a polypeptide of the invention contains a secretory signal sequence that facilitates secretion of the protein. Typically, the signal sequence is positioned in the coding region of the polynucleotide to be expressed towards or at the 5' end of the coding region. The signal sequence may be 10 homologous or heterologous to the polynucleotide of interest and may be homologous or heterologous to the cells to be transfected. Additionally, the signal sequence may be chemically synthesized using methods known in the art.

Any mode of administration of any of the above-described polynucleotides constructs can be used so long as the mode results in the expression of one or more 15 molecules in an amount sufficient to provide a therapeutic effect. This includes direct needle injection, systemic injection, catheter infusion, biolistic injectors, particle accelerators (i.e., "gene guns"), gelfoam sponge depots, other commercially available depot materials, osmotic pumps (e.g., Alza minipumps), oral or suppository solid (tablet or pill) pharmaceutical formulations, and decanting or topical applications 20 during surgery. For example, direct injection of naked calcium phosphate-precipitated plasmid into rat liver and rat spleen or a protein-coated plasmid into the portal vein has resulted in gene expression of the foreign gene in the rat livers. (Kaneda et al., Science, 243:375 (1989)).

A preferred method of local administration is by direct injection. Preferably, a 25 recombinant molecule of the present invention complexed with a delivery vehicle is administered by direct injection into or locally within the area of arteries. Administration of a composition locally within the area of arteries refers to injecting the composition centimeters and preferably, millimeters within arteries.

Another method of local administration is to contact a polynucleotide 30 construct of the present invention in or around a surgical wound. For example, a patient can undergo surgery and the polynucleotide construct can be coated on the

surface of tissue inside the wound or the construct can be injected into areas of tissue inside the wound.

Therapeutic compositions useful in systemic administration, include recombinant molecules of the present invention complexed to a targeted delivery vehicle of the present invention. Suitable delivery vehicles for use with systemic administration comprise liposomes comprising ligands for targeting the vehicle to a particular site.

Preferred methods of systemic administration, include intravenous injection, aerosol, oral and percutaneous (topical) delivery. Intravenous injections can be performed using methods standard in the art. Aerosol delivery can also be performed using methods standard in the art (see, for example, Stribling et al., Proc. Natl. Acad. Sci. USA, 189:11277-11281 (1992), which is incorporated herein by reference). Oral delivery can be performed by complexing a polynucleotide construct of the present invention to a carrier capable of withstanding degradation by digestive enzymes in the gut of an animal. Examples of such carriers, include plastic capsules or tablets, such as those known in the art. Topical delivery can be performed by mixing a polynucleotide construct of the present invention with a lipophilic reagent (e.g., DMSO) that is capable of passing into the skin.

Determining an effective amount of substance to be delivered can depend upon a number of factors including, for example, the chemical structure and biological activity of the substance, the age and weight of the animal, the precise condition requiring treatment and its severity, and the route of administration. The frequency of treatments depends upon a number of factors, such as the amount of polynucleotide constructs administered per dose, as well as the health and history of the subject. The precise amount, number of doses, and timing of doses will be determined by the attending physician or veterinarian. Therapeutic compositions of the present invention can be administered to any animal, preferably to mammals and birds. Preferred mammals include humans, dogs, cats, mice, rats, rabbits sheep, cattle, horses and pigs, with humans being particularly preferred.

#### Biological Activities

The polynucleotides or polypeptides, or agonists or antagonists of the present invention can be used in assays to test for one or more biological activities. If these

polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides or polypeptides, or agonists or antagonists could be used to treat the associated disease.

5

#### Immune Activity

The polynucleotides or polypeptides, or agonists or antagonists of the present invention may be useful in treating, preventing, and/or diagnosing diseases, disorders, and/or conditions of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop  
 10 through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune diseases, disorders, and/or conditions may be genetic, somatic, such as cancer or some autoimmune diseases, disorders, and/or conditions, acquired (e.g., by chemotherapy or toxins), or infectious.  
 15 Moreover, a polynucleotides or polypeptides, or agonists or antagonists of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotides or polypeptides, or agonists or antagonists of the present invention may be useful in treating, preventing, and/or diagnosing diseases, disorders,  
 20 and/or conditions of hematopoietic cells. A polynucleotides or polypeptides, or agonists or antagonists of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat or prevent those diseases, disorders, and/or conditions associated with a decrease in certain (or many) types hematopoietic cells. Examples  
 25 of immunologic deficiency syndromes include, but are not limited to: blood protein diseases, disorders, and/or conditions (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe  
 30 combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polynucleotides or polypeptides, or agonists or antagonists of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotides or polypeptides, or agonists or antagonists of the present invention could be used to treat or prevent blood coagulation diseases, disorders, and/or conditions (e.g., afibrinogenemia, factor deficiencies, arterial thrombosis, venous thrombosis, etc.), blood platelet diseases, disorders, and/or conditions (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotides or polypeptides, or agonists or antagonists of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. Polynucleotides or polypeptides, or agonists or antagonists of the present invention are may also be useful for the detection, prognosis, treatment, and/or prevention of heart attacks (infarction), strokes, scarring, fibrinolysis, uncontrolled bleeding, uncontrolled coagulation, uncontrolled complement fixation, and/or inflammation.

A polynucleotides or polypeptides, or agonists or antagonists of the present invention may also be useful in treating, preventing, and/or diagnosing autoimmune diseases, disorders, and/or conditions. Many autoimmune diseases, disorders, and/or conditions result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polynucleotides or polypeptides, or agonists or antagonists of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune diseases, disorders, and/or conditions.

Examples of autoimmune diseases, disorders, and/or conditions that can be treated, prevented, and/or diagnosed or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic

Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

5 Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated, prevented, and/or diagnosed by polynucleotides or polypeptides, or agonists or antagonists of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

10 A polynucleotides or polypeptides, or agonists or antagonists of the present invention may also be used to treat, prevent, and/or diagnose organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polynucleotides or polypeptides, 15 or agonists or antagonists of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

20 Similarly, a polynucleotides or polypeptides, or agonists or antagonists of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide or agonists or antagonist may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat, prevent, and/or diagnose inflammatory conditions, both chronic and acute conditions, including chronic prostatitis, granulomatous prostatitis and malacoplakia, inflammation associated with infection (e.g., septic shock, sepsis, or 25 systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

#### Hyperproliferative Disorders

30 A polynucleotides or polypeptides, or agonists or antagonists of the invention can be used to treat, prevent, and/or diagnose hyperproliferative diseases, disorders, and/or conditions, including neoplasms. A polynucleotides or polypeptides, or

agonists or antagonists of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polynucleotides or polypeptides, or agonists or antagonists of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

5           For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative diseases, disorders, and/or conditions can be treated, prevented, and/or diagnosed. This immune response may be increased by either enhancing an existing immune response, or by initiating a new  
10 immune response. Alternatively, decreasing an immune response may also be a method of treating, preventing, and/or diagnosing hyperproliferative diseases, disorders, and/or conditions, such as a chemotherapeutic agent.

          Examples of hyperproliferative diseases, disorders, and/or conditions that can be treated, prevented, and/or diagnosed by polynucleotides or polypeptides, or  
15 agonists or antagonists of the present invention include, but are not limited to neoplasms located in the: colon, abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

20           Similarly, other hyperproliferative diseases, disorders, and/or conditions can also be treated, prevented, and/or diagnosed by a polynucleotides or polypeptides, or agonists or antagonists of the present invention. Examples of such hyperproliferative diseases, disorders, and/or conditions include, but are not limited to: hypergammaglobulinemia, lymphoproliferative diseases, disorders, and/or conditions,  
25 paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstrom's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

          One preferred embodiment utilizes polynucleotides of the present invention to inhibit aberrant cellular division, by gene therapy using the present invention, and/or  
30 protein fusions or fragments thereof.

          Thus, the present invention provides a method for treating or preventing cell proliferative diseases, disorders, and/or conditions by inserting into an abnormally

proliferating cell a polynucleotide of the present invention, wherein said polynucleotide represses said expression.

Another embodiment of the present invention provides a method of treating or preventing cell-proliferative diseases, disorders, and/or conditions in individuals comprising administration of one or more active gene copies of the present invention to an abnormally proliferating cell or cells. In a preferred embodiment, polynucleotides of the present invention is a DNA construct comprising a recombinant expression vector effective in expressing a DNA sequence encoding said polynucleotides. In another preferred embodiment of the present invention, the DNA construct encoding the polynucleotides of the present invention is inserted into cells to be treated utilizing a retrovirus, or more Preferably an adenoviral vector (See G J. Nabel, et. al., PNAS 1999 96: 324-326, which is hereby incorporated by reference). In a most preferred embodiment, the viral vector is defective and will not transform non-proliferating cells, only proliferating cells. Moreover, in a preferred embodiment, the polynucleotides of the present invention inserted into proliferating cells either alone, or in combination with or fused to other polynucleotides, can then be modulated via an external stimulus (i.e. magnetic, specific small molecule, chemical, or drug administration, etc.), which acts upon the promoter upstream of said polynucleotides to induce expression of the encoded protein product. As such the beneficial therapeutic affect of the present invention may be expressly modulated (i.e. to increase, decrease, or inhibit expression of the present invention) based upon said external stimulus.

Polynucleotides of the present invention may be useful in repressing expression of oncogenic genes or antigens. By "repressing expression of the oncogenic genes " is intended the suppression of the transcription of the gene, the degradation of the gene transcript (pre-message RNA), the inhibition of splicing, the destruction of the messenger RNA, the prevention of the post-translational modifications of the protein, the destruction of the protein, or the inhibition of the normal function of the protein.

For local administration to abnormally proliferating cells, polynucleotides of the present invention may be administered by any method known to those of skill in the art including, but not limited to transfection, electroporation, microinjection of

cells, or in vehicles such as liposomes, lipofectin, or as naked polynucleotides, or any other method described throughout the specification. The polynucleotide of the present invention may be delivered by known gene delivery systems such as, but not limited to, retroviral vectors (Gilboa, J. Virology 44:845 (1982); Hocke, Nature 320:275 (1986); Wilson, et al., Proc. Natl. Acad. Sci. U.S.A. 85:3014), vaccinia virus system (Chakrabarty et al., Mol. Cell Biol. 5:3403 (1985) or other efficient DNA delivery systems (Yates et al., Nature 313:812 (1985)) known to those skilled in the art. These references are exemplary only and are hereby incorporated by reference. In order to specifically deliver or transfect cells which are abnormally proliferating and spare non-dividing cells, it is preferable to utilize a retrovirus, or adenoviral (as described in the art and elsewhere herein) delivery system known to those of skill in the art. Since host DNA replication is required for retroviral DNA to integrate and the retrovirus will be unable to self replicate due to the lack of the retrovirus genes needed for its life cycle. Utilizing such a retroviral delivery system for polynucleotides of the present invention will target said gene and constructs to abnormally proliferating cells and will spare the non-dividing normal cells.

The polynucleotides of the present invention may be delivered directly to cell proliferative disorder/disease sites in internal organs, body cavities and the like by use of imaging devices used to guide an injecting needle directly to the disease site. The polynucleotides of the present invention may also be administered to disease sites at the time of surgical intervention.

By "cell proliferative disease" is meant any human or animal disease or disorder, affecting any one or any combination of organs, cavities, or body parts, which is characterized by single or multiple local abnormal proliferations of cells, groups of cells, or tissues, whether benign or malignant.

Any amount of the polynucleotides of the present invention may be administered as long as it has a biologically inhibiting effect on the proliferation of the treated cells. Moreover, it is possible to administer more than one of the polynucleotide of the present invention simultaneously to the same site. By "biologically inhibiting" is meant partial or total growth inhibition as well as decreases in the rate of proliferation or growth of the cells. The biologically inhibitory dose may be determined by assessing the effects of the polynucleotides of the present

invention on target malignant or abnormally proliferating cell growth in tissue culture, tumor growth in animals and cell cultures, or any other method known to one of ordinary skill in the art.

5 The present invention is further directed to antibody-based therapies which involve administering of anti-polypeptides and anti-polynucleotide antibodies to a mammalian, preferably human, patient for treating, preventing, and/or diagnosing one or more of the described diseases, disorders, and/or conditions. Methods for producing anti-polypeptides and anti-polynucleotide antibodies polyclonal and monoclonal antibodies are described in detail elsewhere herein. Such antibodies may  
10 be provided in pharmaceutically acceptable compositions as known in the art or as described herein.

A summary of the ways in which the antibodies of the present invention may be used therapeutically includes binding polynucleotides or polypeptides of the present invention locally or systemically in the body or by direct cytotoxicity of the  
15 antibody, e.g. as mediated by complement (CDC) or by effector cells (ADCC). Some of these approaches are described in more detail below. Armed with the teachings provided herein, one of ordinary skill in the art will know how to use the antibodies of the present invention for diagnostic, monitoring or therapeutic purposes without undue experimentation.

20 In particular, the antibodies, fragments and derivatives of the present invention are useful for treating, preventing, and/or diagnosing a subject having or developing cell proliferative and/or differentiation diseases, disorders, and/or conditions as described herein. Such treatment comprises administering a single or multiple doses of the antibody, or a fragment, derivative, or a conjugate thereof.

25 The antibodies of this invention may be advantageously utilized in combination with other monoclonal or chimeric antibodies, or with lymphokines or hematopoietic growth factors, for example, which serve to increase the number or activity of effector cells which interact with the antibodies.

30 It is preferred to use high affinity and/or potent in vivo inhibiting and/or neutralizing antibodies against polypeptides or polynucleotides of the present invention, fragments or regions thereof, for both immunoassays directed to and therapy of diseases, disorders, and/or conditions related to polynucleotides or

polypeptides, including fragments thereof, of the present invention. Such antibodies, fragments, or regions, will preferably have an affinity for polynucleotides or polypeptides, including fragments thereof. Preferred binding affinities include those with a dissociation constant or  $K_d$  less than  $5 \times 10^{-6}M$ ,  $10^{-6}M$ ,  $5 \times 10^{-7}M$ ,  $10^{-7}M$ ,  
5  $5 \times 10^{-8}M$ ,  $10^{-8}M$ ,  $5 \times 10^{-9}M$ ,  $10^{-9}M$ ,  $5 \times 10^{-10}M$ ,  $10^{-10}M$ ,  $5 \times 10^{-11}M$ ,  $10^{-11}M$ ,  $5 \times 10^{-12}M$ ,  $10^{-12}M$ ,  $5 \times 10^{-13}M$ ,  $10^{-13}M$ ,  $5 \times 10^{-14}M$ ,  $10^{-14}M$ ,  $5 \times 10^{-15}M$ , and  $10^{-15}M$ .

Moreover, polypeptides of the present invention may be useful in inhibiting the angiogenesis of proliferative cells or tissues, either alone, as a protein fusion, or in  
10 combination with other polypeptides directly or indirectly, as described elsewhere herein. In a most preferred embodiment, said anti-angiogenesis effect may be achieved indirectly, for example, through the inhibition of hematopoietic, tumor-specific cells, such as tumor-associated macrophages (See Joseph IB, et al. J Natl Cancer Inst, 90(21):1648-53 (1998), which is hereby incorporated by reference).  
15 Antibodies directed to polypeptides or polynucleotides of the present invention may also result in inhibition of angiogenesis directly, or indirectly (See Witte L, et al., Cancer Metastasis Rev. 17(2):155-61 (1998), which is hereby incorporated by reference)).

Polypeptides, including protein fusions, of the present invention, or fragments  
20 thereof may be useful in inhibiting proliferative cells or tissues through the induction of apoptosis. Said polypeptides may act either directly, or indirectly to induce apoptosis of proliferative cells and tissues, for example in the activation of a death-domain receptor, such as tumor necrosis factor (TNF) receptor-1, CD95 (Fas/APO-1), TNF-receptor-related apoptosis-mediated protein (TRAMP) and TNF-related  
25 apoptosis-inducing ligand (TRAIL) receptor-1 and -2 (See Schulze-Osthoff K, et al., Eur J Biochem 254(3):439-59 (1998), which is hereby incorporated by reference). Moreover, in another preferred embodiment of the present invention, said polypeptides may induce apoptosis through other mechanisms, such as in the activation of other proteins which will activate apoptosis, or through stimulating the  
30 expression of said proteins, either alone or in combination with small molecule drugs or adjuvants, such as apoptonin, galectins, thioredoxins, antiinflammatory proteins (See for example, Mutat. Res. 400(1-2):447-55 (1998), Med Hypotheses.50(5):423-33

(1998), Chem. Biol. Interact. Apr 24;111-112:23-34 (1998), J Mol Med.76(6):402-12 (1998), Int. J. Tissue React. 20(1):3-15 (1998), which are all hereby incorporated by reference).

Polypeptides, including protein fusions to, or fragments thereof, of the present invention are useful in inhibiting the metastasis of proliferative cells or tissues. Inhibition may occur as a direct result of administering polypeptides, or antibodies directed to said polypeptides as described elsewhere herein, or indirectly, such as activating the expression of proteins known to inhibit metastasis, for example alpha 4 integrins, (See, e.g., Curr Top Microbiol Immunol 1998;231:125-41, which is hereby incorporated by reference). Such therapeutic affects of the present invention may be achieved either alone, or in combination with small molecule drugs or adjuvants.

In another embodiment, the invention provides a method of delivering compositions containing the polypeptides of the invention (e.g., compositions containing polypeptides or polypeptide antibodies associated with heterologous polypeptides, heterologous nucleic acids, toxins, or prodrugs) to targeted cells expressing the polypeptide of the present invention. Polypeptides or polypeptide antibodies of the invention may be associated with heterologous polypeptides, heterologous nucleic acids, toxins, or prodrugs via hydrophobic, hydrophilic, ionic and/or covalent interactions.

Polypeptides, protein fusions to, or fragments thereof, of the present invention are useful in enhancing the immunogenicity and/or antigenicity of proliferating cells or tissues, either directly, such as would occur if the polypeptides of the present invention 'vaccinated' the immune response to respond to proliferative antigens and immunogens, or indirectly, such as in activating the expression of proteins known to enhance the immune response (e.g. chemokines), to said antigens and immunogens.

#### Cardiovascular Disorders

Polynucleotides or polypeptides, or agonists or antagonists of the invention may be used to treat, prevent, and/or diagnose cardiovascular diseases, disorders, and/or conditions, including peripheral artery disease, such as limb ischemia.

Cardiovascular diseases, disorders, and/or conditions include cardiovascular abnormalities, such as arterio-arterial fistula, arteriovenous fistula, cerebral arteriovenous malformations, congenital heart defects, pulmonary atresia, and

Scimitar Syndrome. Congenital heart defects include aortic coarctation, cor triatriatum, coronary vessel anomalies, crisscross heart, dextrocardia, patent ductus arteriosus, Ebstein's anomaly, Eisenmenger complex, hypoplastic left heart syndrome, levocardia, tetralogy of fallot, transposition of great vessels, double outlet right ventricle, tricuspid atresia, persistent truncus arteriosus, and heart septal defects, such as aortopulmonary septal defect, endocardial cushion defects, Lutembacher's Syndrome, trilog of Fallot, ventricular heart septal defects.

Cardiovascular diseases, disorders, and/or conditions also include heart disease, such as arrhythmias, carcinoid heart disease, high cardiac output, low cardiac output, cardiac tamponade, endocarditis (including bacterial), heart aneurysm, cardiac arrest, congestive heart failure, congestive cardiomyopathy, paroxysmal dyspnea, cardiac edema, heart hypertrophy, congestive cardiomyopathy, left ventricular hypertrophy, right ventricular hypertrophy, post-infarction heart rupture, ventricular septal rupture, heart valve diseases, myocardial diseases, myocardial ischemia, pericardial effusion, pericarditis (including constrictive and tuberculous), pneumopericardium, postpericardiotomy syndrome, pulmonary heart disease, rheumatic heart disease, ventricular dysfunction, hyperemia, cardiovascular pregnancy complications, Scimitar Syndrome, cardiovascular syphilis, and cardiovascular tuberculosis.

Arrhythmias include sinus arrhythmia, atrial fibrillation, atrial flutter, bradycardia, extrasystole, Adams-Stokes Syndrome, bundle-branch block, sinoatrial block, long QT syndrome, parasystole, Lown-Ganong-Levine Syndrome, Mahaim-type pre-excitation syndrome, Wolff-Parkinson-White syndrome, sick sinus syndrome, tachycardias, and ventricular fibrillation. Tachycardias include paroxysmal tachycardia, supraventricular tachycardia, accelerated idioventricular rhythm, atrioventricular nodal reentry tachycardia, ectopic atrial tachycardia, ectopic junctional tachycardia, sinoatrial nodal reentry tachycardia, sinus tachycardia, Torsades de Pointes, and ventricular tachycardia.

Heart valve disease include aortic valve insufficiency, aortic valve stenosis, hear murmurs, aortic valve prolapse, mitral valve prolapse, tricuspid valve prolapse, mitral valve insufficiency, mitral valve stenosis, pulmonary atresia, pulmonary valve

insufficiency, pulmonary valve stenosis, tricuspid atresia, tricuspid valve insufficiency, and tricuspid valve stenosis.

Myocardial diseases include alcoholic cardiomyopathy, congestive cardiomyopathy, hypertrophic cardiomyopathy, aortic subvalvular stenosis, pulmonary subvalvular stenosis, restrictive cardiomyopathy, Chagas cardiomyopathy, endocardial fibroelastosis, endomyocardial fibrosis, Kearns Syndrome, myocardial reperfusion injury, and myocarditis.

Myocardial ischemias include coronary disease, such as angina pectoris, coronary aneurysm, coronary arteriosclerosis, coronary thrombosis, coronary vasospasm, myocardial infarction and myocardial stunning.

Cardiovascular diseases also include vascular diseases such as aneurysms, angiodyplasia, angiomas, bacillary angiomas, Hippel-Lindau Disease, Klippel-Trenaunay-Weber Syndrome, Sturge-Weber Syndrome, angioneurotic edema, aortic diseases, Takayasu's Arteritis, aortitis, Leriche's Syndrome, arterial occlusive diseases, arteritis, enarteritis, polyarteritis nodosa, cerebrovascular diseases, disorders, and/or conditions, diabetic angiopathies, diabetic retinopathy, embolisms, thrombosis, erythromelalgia, hemorrhoids, hepatic veno-occlusive disease, hypertension, hypotension, ischemia, peripheral vascular diseases, phlebitis, pulmonary veno-occlusive disease, Raynaud's disease, CREST syndrome, retinal vein occlusion, Scimitar syndrome, superior vena cava syndrome, telangiectasia, ataxia telangiectasia, hereditary hemorrhagic telangiectasia, varicocele, varicose veins, varicose ulcer, vasculitis, and venous insufficiency.

Aneurysms include dissecting aneurysms, false aneurysms, infected aneurysms, ruptured aneurysms, aortic aneurysms, cerebral aneurysms, coronary aneurysms, heart aneurysms, and iliac aneurysms.

Arterial occlusive diseases include arteriosclerosis, intermittent claudication, carotid stenosis, fibromuscular dysplasias, mesenteric vascular occlusion, Moyamoya disease, renal artery obstruction, retinal artery occlusion, and thromboangiitis obliterans.

Cerebrovascular diseases, disorders, and/or conditions include carotid artery diseases, cerebral amyloid angiopathy, cerebral aneurysm, cerebral anoxia, cerebral arteriosclerosis, cerebral arteriovenous malformation, cerebral artery diseases,

cerebral embolism and thrombosis, carotid artery thrombosis, sinus thrombosis, Wallenberg's syndrome, cerebral hemorrhage, epidural hematoma, subdural hematoma, subarachnoid hemorrhage, cerebral infarction, cerebral ischemia (including transient), subclavian steal syndrome, periventricular leukomalacia, vascular headache, cluster headache, migraine, and vertebrobasilar insufficiency.

Embolisms include air embolisms, amniotic fluid embolisms, cholesterol embolisms, blue toe syndrome, fat embolisms, pulmonary embolisms, and thromboembolisms. Thrombosis include coronary thrombosis, hepatic vein thrombosis, retinal vein occlusion, carotid artery thrombosis, sinus thrombosis, Wallenberg's syndrome, and thrombophlebitis.

Ischemia includes cerebral ischemia, ischemic colitis, compartment syndromes, anterior compartment syndrome, myocardial ischemia, reperfusion injuries, and peripheral limb ischemia. Vasculitis includes aortitis, arteritis, Behcet's Syndrome, Churg-Strauss Syndrome, mucocutaneous lymph node syndrome, thromboangiitis obliterans, hypersensitivity vasculitis, Schoenlein-Henoch purpura, allergic cutaneous vasculitis, and Wegener's granulomatosis.

Polynucleotides or polypeptides, or agonists or antagonists of the invention, are especially effective for the treatment of critical limb ischemia and coronary disease.

Polypeptides may be administered using any method known in the art, including, but not limited to, direct needle injection at the delivery site, intravenous injection, topical administration, catheter infusion, biolistic injectors, particle accelerators, gelfoam sponge depots, other commercially available depot materials, osmotic pumps, oral or suppository solid pharmaceutical formulations, decanting or topical applications during surgery, aerosol delivery. Such methods are known in the art. Polypeptides of the invention may be administered as part of a Therapeutic, described in more detail below. Methods of delivering polynucleotides of the invention are described in more detail herein.

#### Anti-Angiogenesis Activity

The naturally occurring balance between endogenous stimulators and inhibitors of angiogenesis is one in which inhibitory influences predominate. Rastinejad et al., Cell 56:345-355 (1989). In those rare instances in which

neovascularization occurs under normal physiological conditions, such as wound healing, organ regeneration, embryonic development, and female reproductive processes, angiogenesis is stringently regulated and spatially and temporally delimited. Under conditions of pathological angiogenesis such as that characterizing

5 solid tumor growth, these regulatory controls fail. Unregulated angiogenesis becomes pathologic and sustains progression of many neoplastic and non-neoplastic diseases. A number of serious diseases are dominated by abnormal neovascularization including solid tumor growth and metastases, arthritis, some types of eye diseases, disorders, and/or conditions, and psoriasis. See, e.g., reviews by Moses et al., *Biotech.*

10 9:630-634 (1991); Folkman et al., *N. Engl. J. Med.*, 333:1757-1763 (1995); Auerbach et al., *J. Microvasc. Res.* 29:401-411 (1985); Folkman, *Advances in Cancer Research*, eds. Klein and Weinhouse, Academic Press, New York, pp. 175-203 (1985); Patz, *Am. J. Ophthalmol.* 94:715-743 (1982); and Folkman et al., *Science* 221:719-725 (1983). In a number of pathological conditions, the process of angiogenesis

15 contributes to the disease state. For example, significant data have accumulated which suggest that the growth of solid tumors is dependent on angiogenesis. Folkman and Klagsbrun, *Science* 235:442-447 (1987).

The present invention provides for treatment of diseases, disorders, and/or conditions associated with neovascularization by administration of the

20 polynucleotides and/or polypeptides of the invention, as well as agonists or antagonists of the present invention. Malignant and metastatic conditions which can be treated with the polynucleotides and polypeptides, or agonists or antagonists of the invention include, but are not limited to, malignancies, solid tumors, and cancers described herein and otherwise known in the art (for a review of such disorders, see

25 Fishman et al., *Medicine*, 2d Ed., J. B. Lippincott Co., Philadelphia (1985)). Thus, the present invention provides a method of treating, preventing, and/or diagnosing an angiogenesis-related disease and/or disorder, comprising administering to an individual in need thereof a therapeutically effective amount of a polynucleotide, polypeptide, antagonist and/or agonist of the invention. For example, polynucleotides,

30 polypeptides, antagonists and/or agonists may be utilized in a variety of additional methods in order to therapeutically treat or prevent a cancer or tumor. Cancers which may be treated, prevented, and/or diagnosed with polynucleotides, polypeptides,

antagonists and/or agonists include, but are not limited to solid tumors, including prostate, lung, breast, ovarian, stomach, pancreas, larynx, esophagus, testes, liver, parotid, biliary tract, colon, rectum, cervix, uterus, endometrium, kidney, bladder, thyroid cancer; primary tumors and metastases; melanomas; glioblastoma; Kaposi's sarcoma; leiomyosarcoma; non- small cell lung cancer; colorectal cancer; advanced malignancies; and blood born tumors such as leukemias. For example, polynucleotides, polypeptides, antagonists and/or agonists may be delivered topically, in order to treat or prevent cancers such as skin cancer, head and neck tumors, breast tumors, and Kaposi's sarcoma.

Within yet other aspects, polynucleotides, polypeptides, antagonists and/or agonists may be utilized to treat superficial forms of bladder cancer by, for example, intravesical administration. Polynucleotides, polypeptides, antagonists and/or agonists may be delivered directly into the tumor, or near the tumor site, via injection or a catheter. Of course, as the artisan of ordinary skill will appreciate, the appropriate mode of administration will vary according to the cancer to be treated. Other modes of delivery are discussed herein.

Polynucleotides, polypeptides, antagonists and/or agonists may be useful in treating, preventing, and/or diagnosing other diseases, disorders, and/or conditions, besides cancers, which involve angiogenesis. These diseases, disorders, and/or conditions include, but are not limited to: benign tumors, for example hemangiomas, acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas; arteriosclerotic plaques; ocular angiogenic diseases, for example, diabetic retinopathy, retinopathy of prematurity, macular degeneration, corneal graft rejection, neovascular glaucoma, retrolental fibroplasia, rubeosis, retinoblastoma, uveitis and Pterygia (abnormal blood vessel growth) of the eye; rheumatoid arthritis; psoriasis; delayed wound healing; endometriosis; vasculogenesis; granulations; hypertrophic scars (keloids); nonunion fractures; scleroderma; trachoma; vascular adhesions; myocardial angiogenesis; coronary collaterals; cerebral collaterals; arteriovenous malformations; ischemic limb angiogenesis; Osler-Webber Syndrome; plaque neovascularization; telangiectasia; hemophilic joints; angiofibroma; fibromuscular dysplasia; wound granulation; Crohn's disease; and atherosclerosis.

For example, within one aspect of the present invention methods are provided for treating, preventing, and/or diagnosing hypertrophic scars and keloids, comprising the step of administering a polynucleotide, polypeptide, antagonist and/or agonist of the invention to a hypertrophic scar or keloid.

5           Within one embodiment of the present invention polynucleotides, polypeptides, antagonists and/or agonists are directly injected into a hypertrophic scar or keloid, in order to prevent the progression of these lesions. This therapy is of particular value in the prophylactic treatment of conditions which are known to result in the development of hypertrophic scars and keloids (e.g., burns), and is preferably  
10 initiated after the proliferative phase has had time to progress (approximately 14 days after the initial injury), but before hypertrophic scar or keloid development. As noted above, the present invention also provides methods for treating, preventing, and/or diagnosing neovascular diseases of the eye, including for example, corneal neovascularization, neovascular glaucoma, proliferative diabetic retinopathy,  
15 retrolental fibroplasia and macular degeneration.

Moreover, Ocular diseases, disorders, and/or conditions associated with neovascularization which can be treated, prevented, and/or diagnosed with the polynucleotides and polypeptides of the present invention (including agonists and/or antagonists) include, but are not limited to: neovascular glaucoma, diabetic  
20 retinopathy, retinoblastoma, retrolental fibroplasia, uveitis, retinopathy of prematurity macular degeneration, corneal graft neovascularization, as well as other eye inflammatory diseases, ocular tumors and diseases associated with choroidal or iris neovascularization. See, e.g., reviews by Waltman et al., *Am. J. Ophthalm.* 85:704-710 (1978) and Gartner et al., *Surv. Ophthalm.* 22:291-312 (1978).

25           Thus, within one aspect of the present invention methods are provided for treating or preventing neovascular diseases of the eye such as corneal neovascularization (including corneal graft neovascularization), comprising the step of administering to a patient a therapeutically effective amount of a compound (as described above) to the cornea, such that the formation of blood vessels is inhibited.  
30 Briefly, the cornea is a tissue which normally lacks blood vessels. In certain pathological conditions however, capillaries may extend into the cornea from the pericorneal vascular plexus of the limbus. When the cornea becomes vascularized, it

also becomes clouded, resulting in a decline in the patient's visual acuity. Visual loss may become complete if the cornea completely opacitates. A wide variety of diseases, disorders, and/or conditions can result in corneal neovascularization, including for example, corneal infections (e.g., trachoma, herpes simplex keratitis, leishmaniasis and onchocerciasis), immunological processes (e.g., graft rejection and Stevens-Johnson's syndrome), alkali burns, trauma, inflammation (of any cause), toxic and nutritional deficiency states, and as a complication of wearing contact lenses.

Within particularly preferred embodiments of the invention, may be prepared for topical administration in saline (combined with any of the preservatives and antimicrobial agents commonly used in ocular preparations), and administered in eyedrop form. The solution or suspension may be prepared in its pure form and administered several times daily. Alternatively, anti-angiogenic compositions, prepared as described above, may also be administered directly to the cornea. Within preferred embodiments, the anti-angiogenic composition is prepared with a muco-adhesive polymer which binds to cornea. Within further embodiments, the anti-angiogenic factors or anti-angiogenic compositions may be utilized as an adjunct to conventional steroid therapy. Topical therapy may also be useful prophylactically in corneal lesions which are known to have a high probability of inducing an angiogenic response (such as chemical burns). In these instances the treatment, likely in combination with steroids, may be instituted immediately to help prevent subsequent complications.

Within other embodiments, the compounds described above may be injected directly into the corneal stroma by an ophthalmologist under microscopic guidance. The preferred site of injection may vary with the morphology of the individual lesion, but the goal of the administration would be to place the composition at the advancing front of the vasculature (i.e., interspersed between the blood vessels and the normal cornea). In most cases this would involve perilimbic corneal injection to "protect" the cornea from the advancing blood vessels. This method may also be utilized shortly after a corneal insult in order to prophylactically prevent corneal neovascularization. In this situation the material could be injected in the perilimbic cornea interspersed between the corneal lesion and its undesired potential limbic blood supply. Such methods may also be utilized in a similar fashion to prevent capillary invasion of

transplanted corneas. In a sustained-release form injections might only be required 2-3 times per year. A steroid could also be added to the injection solution to reduce inflammation resulting from the injection itself.

5        Within another aspect of the present invention, methods are provided for treating or preventing neovascular glaucoma, comprising the step of administering to a patient a therapeutically effective amount of a polynucleotide, polypeptide, antagonist and/or agonist to the eye, such that the formation of blood vessels is inhibited. In one embodiment, the compound may be administered topically to the eye in order to treat or prevent early forms of neovascular glaucoma. Within other  
10        embodiments, the compound may be implanted by injection into the region of the anterior chamber angle. Within other embodiments, the compound may also be placed in any location such that the compound is continuously released into the aqueous humor. Within another aspect of the present invention, methods are provided for treating or preventing proliferative diabetic retinopathy, comprising the step of  
15        administering to a patient a therapeutically effective amount of a polynucleotide, polypeptide, antagonist and/or agonist to the eyes, such that the formation of blood vessels is inhibited.

      Within particularly preferred embodiments of the invention, proliferative diabetic retinopathy may be treated by injection into the aqueous humor or the  
20        vitreous, in order to increase the local concentration of the polynucleotide, polypeptide, antagonist and/or agonist in the retina. Preferably, this treatment should be initiated prior to the acquisition of severe disease requiring photocoagulation.

      Within another aspect of the present invention, methods are provided for treating or preventing retrolental fibroplasia, comprising the step of administering to a  
25        patient a therapeutically effective amount of a polynucleotide, polypeptide, antagonist and/or agonist to the eye, such that the formation of blood vessels is inhibited. The compound may be administered topically, via intravitreal injection and/or via intraocular implants.

      Additionally, diseases, disorders, and/or conditions which can be treated, prevented, and/or diagnosed with the polynucleotides, polypeptides, agonists and/or  
30        agonists include, but are not limited to, hemangioma, arthritis, psoriasis, angiofibroma, atherosclerotic plaques, delayed wound healing, granulations,

hemophilic joints, hypertrophic scars, nonunion fractures, Osler-Weber syndrome, pyogenic granuloma, scleroderma, trachoma, and vascular adhesions.

Moreover, diseases, disorders, and/or conditions and/or states, which can be treated, prevented, and/or diagnosed with the polynucleotides, polypeptides, agonists and/or agonists include, but are not limited to, solid tumors, blood born tumors such as leukemias, tumor metastasis, Kaposi's sarcoma, benign tumors, for example hemangiomas, acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases, for example, diabetic retinopathy, retinopathy of prematurity, macular degeneration, corneal graft rejection, neovascular glaucoma, retrolental fibroplasia, rubeosis, retinoblastoma, and uveitis, delayed wound healing, endometriosis, vasculogenesis, granulations, hypertrophic scars (keloids), nonunion fractures, scleroderma, trachoma, vascular adhesions, myocardial angiogenesis, coronary collaterals, cerebral collaterals, arteriovenous malformations, ischemic limb angiogenesis, Osler-Webber Syndrome, plaque neovascularization, telangiectasia, hemophilic joints, angiofibroma fibromuscular dysplasia, wound granulation, Crohn's disease, atherosclerosis, birth control agent by preventing vascularization required for embryo implantation controlling menstruation, diseases that have angiogenesis as a pathologic consequence such as cat scratch disease (Rochela minalia quintosa), ulcers (*Helicobacter pylori*), Bartonellosis and bacillary angiomatosis.

In one aspect of the birth control method, an amount of the compound sufficient to block embryo implantation is administered before or after intercourse and fertilization have occurred, thus providing an effective method of birth control, possibly a "morning after" method. Polynucleotides, polypeptides, agonists and/or agonists may also be used in controlling menstruation or administered as either a peritoneal lavage fluid or for peritoneal implantation in the treatment of endometriosis.

Polynucleotides, polypeptides, agonists and/or agonists of the present invention may be incorporated into surgical sutures in order to prevent stitch granulomas.

Polynucleotides, polypeptides, agonists and/or agonists may be utilized in a wide variety of surgical procedures. For example, within one aspect of the present

invention a compositions (in the form of, for example, a spray or film) may be utilized to coat or spray an area prior to removal of a tumor, in order to isolate normal surrounding tissues from malignant tissue, and/or to prevent the spread of disease to surrounding tissues. Within other aspects of the present invention, compositions (e.g.,  
5 in the form of a spray) may be delivered via endoscopic procedures in order to coat tumors, or inhibit angiogenesis in a desired locale. Within yet other aspects of the present invention, surgical meshes which have been coated with anti- angiogenic compositions of the present invention may be utilized in any procedure wherein a surgical mesh might be utilized. For example, within one embodiment of the  
10 invention a surgical mesh laden with an anti-angiogenic composition may be utilized during abdominal cancer resection surgery (e.g., subsequent to colon resection) in order to provide support to the structure, and to release an amount of the anti-angiogenic factor.

Within further aspects of the present invention, methods are provided for  
15 treating tumor excision sites, comprising administering a polynucleotide, polypeptide, agonist and/or agonist to the resection margins of a tumor subsequent to excision, such that the local recurrence of cancer and the formation of new blood vessels at the site is inhibited. Within one embodiment of the invention, the anti-angiogenic compound is administered directly to the tumor excision site (e.g., applied by  
20 swabbing, brushing or otherwise coating the resection margins of the tumor with the anti-angiogenic compound). Alternatively, the anti-angiogenic compounds may be incorporated into known surgical pastes prior to administration. Within particularly preferred embodiments of the invention, the anti-angiogenic compounds are applied after hepatic resections for malignancy, and after neurosurgical operations.

25 Within one aspect of the present invention, polynucleotides, polypeptides, agonists and/or agonists may be administered to the resection margin of a wide variety of tumors, including for example, breast, colon, brain and hepatic tumors. For example, within one embodiment of the invention, anti-angiogenic compounds may be administered to the site of a neurological tumor subsequent to excision, such that  
30 the formation of new blood vessels at the site are inhibited.

The polynucleotides, polypeptides, agonists and/or agonists of the present invention may also be administered along with other anti-angiogenic factors.

Representative examples of other anti-angiogenic factors include: Anti-Invasive Factor, retinoic acid and derivatives thereof, paclitaxel, Suramin, Tissue Inhibitor of Metalloproteinase-1, Tissue Inhibitor of Metalloproteinase-2, Plasminogen Activator Inhibitor-1, Plasminogen Activator Inhibitor-2, and various forms of the lighter "d  
5 group" transition metals.

Lighter "d group" transition metals include, for example, vanadium, molybdenum, tungsten, titanium, niobium, and tantalum species. Such transition metal species may form transition metal complexes. Suitable complexes of the above-mentioned transition metal species include oxo transition metal complexes.

10 Representative examples of vanadium complexes include oxo vanadium complexes such as vanadate and vanadyl complexes. Suitable vanadate complexes include metavanadate and orthovanadate complexes such as, for example, ammonium metavanadate, sodium metavanadate, and sodium orthovanadate. Suitable vanadyl complexes include, for example, vanadyl acetylacetonate and vanadyl sulfate  
15 including vanadyl sulfate hydrates such as vanadyl sulfate mono- and trihydrates.

Representative examples of tungsten and molybdenum complexes also include oxo complexes. Suitable oxo tungsten complexes include tungstate and tungsten oxide complexes. Suitable tungstate complexes include ammonium tungstate, calcium tungstate, sodium tungstate dihydrate, and tungstic acid. Suitable tungsten oxides  
20 include tungsten (IV) oxide and tungsten (VI) oxide. Suitable oxo molybdenum complexes include molybdate, molybdenum oxide, and molybdenyl complexes. Suitable molybdate complexes include ammonium molybdate and its hydrates, sodium molybdate and its hydrates, and potassium molybdate and its hydrates. Suitable molybdenum oxides include molybdenum (VI) oxide, molybdenum (VI)  
25 oxide, and molybdic acid. Suitable molybdenyl complexes include, for example, molybdenyl acetylacetonate. Other suitable tungsten and molybdenum complexes include hydroxo derivatives derived from, for example, glycerol, tartaric acid, and sugars.

A wide variety of other anti-angiogenic factors may also be utilized within the  
30 context of the present invention. Representative examples include platelet factor 4; protamine sulphate; sulphated chitin derivatives (prepared from queen crab shells), (Murata et al., Cancer Res. 51:22-26, 1991); Sulphated Polysaccharide Peptidoglycan

Complex (SP- PG) (the function of this compound may be enhanced by the presence of steroids such as estrogen, and tamoxifen citrate); Staurosporine; modulators of matrix metabolism, including for example, proline analogs, cishydroxyproline, d,L-3,4-dehydroproline, Thiaproline, alpha,alpha-dipyridyl, aminopropionitrile fumarate; 5 4-propyl-5-(4-pyridinyl)-2(3H)-oxazolone; Methotrexate; Mitoxantrone; Heparin; Interferons; 2 Macroglobulin-serum; ChIMP-3 (Pavloff et al., J. Bio. Chem. 267:17321-17326, 1992); Chymostatin (Tomkinson et al., Biochem J. 286:475-480, 1992); Cyclodextrin Tetradeccasulfate; Eponemycin; Camptothecin; Fumagillin (Ingber et al., Nature 348:555-557, 1990); Gold Sodium Thiomalate ("GST"; 10 Matsubara and Ziff, J. Clin. Invest. 79:1440-1446, 1987); anticollagenase-serum; alpha2-antiplasmin (Holmes et al., J. Biol. Chem. 262(4):1659-1664, 1987); Bisantrone (National Cancer Institute); Lobenzarit disodium (N-(2)-carboxyphenyl-4-chloroanthronilic acid disodium or "CCA"; Takeuchi et al., Agents Actions 36:312-316, 1992); Thalidomide; Angostatic steroid; AGM-1470; carboxynaminolmidazole; 15 and metalloproteinase inhibitors such as BB94.

#### Diseases at the Cellular Level

Diseases associated with increased cell survival or the inhibition of apoptosis that could be treated, prevented, and/or diagnosed by the polynucleotides or polypeptides and/or antagonists or agonists of the invention, include cancers (such as 20 follicular lymphomas, carcinomas with p53 mutations, and hormone-dependent tumors, including, but not limited to colon cancer, cardiac tumors, pancreatic cancer, melanoma, retinoblastoma, glioblastoma, lung cancer, intestinal cancer, testicular cancer, stomach cancer, neuroblastoma, myxoma, myoma, lymphoma, endothelioma, osteoblastoma, osteoclastoma, osteosarcoma, chondrosarcoma, adenoma, breast 25 cancer, prostate cancer, Kaposi's sarcoma and ovarian cancer); autoimmune diseases, disorders, and/or conditions (such as, multiple sclerosis, Sjogren's syndrome, Hashimoto's thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease, polymyositis, systemic lupus erythematosus and immune-related glomerulonephritis and rheumatoid arthritis) and viral infections (such as herpes viruses, pox viruses and 30 adenoviruses), inflammation, graft v. host disease, acute graft rejection, and chronic graft rejection. In preferred embodiments, the polynucleotides or polypeptides, and/or

agonists or antagonists of the invention are used to inhibit growth, progression, and/or metastasis of cancers, in particular those listed above.

Additional diseases or conditions associated with increased cell survival that could be treated, prevented or diagnosed by the polynucleotides or polypeptides, or agonists or antagonists of the invention, include, but are not limited to, progression, and/or metastases of malignancies and related disorders such as leukemia (including acute leukemias (e.g., acute lymphocytic leukemia, acute myelocytic leukemia (including myeloblastic, promyelocytic, myelomonocytic, monocytic, and erythroleukemia)) and chronic leukemias (e.g., chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia)), polycythemia vera, lymphomas (e.g., Hodgkin's disease and non-Hodgkin's disease), multiple myeloma, Waldenstrom's macroglobulinemia, heavy chain disease, and solid tumors including, but not limited to, sarcomas and carcinomas such as fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, colon carcinoma, pancreatic cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, choriocarcinoma, seminoma, embryonal carcinoma, Wilm's tumor, cervical cancer, testicular tumor, lung carcinoma, small cell lung carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, menangioma, melanoma, neuroblastoma, and retinoblastoma.

Diseases associated with increased apoptosis that could be treated, prevented, and/or diagnosed by the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, include AIDS; neurodegenerative diseases, disorders, and/or conditions (such as Alzheimer's disease, Parkinson's disease, Amyotrophic lateral sclerosis, Retinitis pigmentosa, Cerebellar degeneration and brain tumor or prior associated disease); autoimmune diseases, disorders, and/or conditions (such as, multiple sclerosis, Sjogren's syndrome, Hashimoto's thyroiditis, biliary cirrhosis,

Behcet's disease, Crohn's disease, polymyositis, systemic lupus erythematosus and immune-related glomerulonephritis and rheumatoid arthritis) myelodysplastic syndromes (such as aplastic anemia), graft v. host disease, ischemic injury (such as that caused by myocardial infarction, stroke and reperfusion injury), liver injury (e.g.,  
 5 hepatitis related liver injury, ischemia/reperfusion injury, cholestasis (bile duct injury) and liver cancer); toxin-induced liver disease (such as that caused by alcohol), septic shock, cachexia and anorexia.

#### Wound Healing and Epithelial Cell Proliferation

In accordance with yet a further aspect of the present invention, there is  
 10 provided a process for utilizing the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, for therapeutic purposes, for example, to stimulate epithelial cell proliferation and basal keratinocytes for the purpose of wound healing, and to stimulate hair follicle production and healing of dermal wounds. Polynucleotides or polypeptides, as well as agonists or antagonists of the invention,  
 15 may be clinically useful in stimulating wound healing including surgical wounds, excisional wounds, deep wounds involving damage of the dermis and epidermis, eye tissue wounds, dental tissue wounds, oral cavity wounds, diabetic ulcers, dermal ulcers, cubitus ulcers, arterial ulcers, venous stasis ulcers, burns resulting from heat exposure or chemicals, and other abnormal wound healing conditions such as uremia,  
 20 malnutrition, vitamin deficiencies and complications associated with systemic treatment with steroids, radiation therapy and antineoplastic drugs and antimetabolites. Polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to promote dermal reestablishment subsequent to dermal loss

The polynucleotides or polypeptides, and/or agonists or antagonists of the  
 25 invention, could be used to increase the adherence of skin grafts to a wound bed and to stimulate re-epithelialization from the wound bed. The following are a non-exhaustive list of grafts that polynucleotides or polypeptides, agonists or antagonists of the invention, could be used to increase adherence to a wound bed: autografts, artificial skin, allografts, autodermic graft, autoepidermic grafts, avacular grafts,  
 30 Blair-Brown grafts, bone graft, brephoplastic grafts, cutis graft, delayed graft, dermic graft, epidermic graft, fascia graft, full thickness graft, heterologous graft, xenograft, homologous graft, hyperplastic graft, lamellar graft, mesh graft, mucosal graft, Ollier-

Thiersch graft, omentoplastic graft, patch graft, pedicle graft, penetrating graft, split skin graft, thick split graft. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, can be used to promote skin strength and to improve the appearance of aged skin.

5           It is believed that the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, will also produce changes in hepatocyte proliferation, and epithelial cell proliferation in the lung, breast, pancreas, stomach, small intestine, and large intestine. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could promote proliferation of epithelial cells such as  
10   sebocytes, hair follicles, hepatocytes, type II pneumocytes, mucin-producing goblet cells, and other epithelial cells and their progenitors contained within the skin, lung, liver, and gastrointestinal tract. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, may promote proliferation of endothelial cells, keratinocytes, and basal keratinocytes.

15           The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could also be used to reduce the side effects of gut toxicity that result from radiation, chemotherapy treatments or viral infections. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, may have a cytoprotective effect on the small intestine mucosa. The polynucleotides or  
20   polypeptides, and/or agonists or antagonists of the invention, may also stimulate healing of mucositis (mouth ulcers) that result from chemotherapy and viral infections.

          The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could further be used in full regeneration of skin in full and partial  
25   thickness skin defects, including burns, (i.e., repopulation of hair follicles, sweat glands, and sebaceous glands), treatment of other skin defects such as psoriasis. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to treat epidermolysis bullosa, a defect in adherence of the epidermis to the underlying dermis which results in frequent, open and painful blisters by accelerating  
30   reepithelialization of these lesions. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could also be used to treat gastric and duodenal ulcers and help heal by scar formation of the mucosal lining and

regeneration of glandular mucosa and duodenal mucosal lining more rapidly. Inflammatory bowel diseases, such as Crohn's disease and ulcerative colitis, are diseases which result in destruction of the mucosal surface of the small or large intestine, respectively. Thus, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to promote the resurfacing of the mucosal surface to aid more rapid healing and to prevent progression of inflammatory bowel disease. Treatment with the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, is expected to have a significant effect on the production of mucus throughout the gastrointestinal tract and could be used to protect the intestinal mucosa from injurious substances that are ingested or following surgery. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to treat diseases associated with the under expression of the polynucleotides of the invention.

Moreover, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to prevent and heal damage to the lungs due to various pathological states. A growth factor such as the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, which could stimulate proliferation and differentiation and promote the repair of alveoli and bronchiolar epithelium to prevent or treat acute or chronic lung damage. For example, emphysema, which results in the progressive loss of alveoli, and inhalation injuries, i.e., resulting from smoke inhalation and burns, that cause necrosis of the bronchiolar epithelium and alveoli could be effectively treated, prevented, and/or diagnosed using the polynucleotides or polypeptides, and/or agonists or antagonists of the invention. Also, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to stimulate the proliferation of and differentiation of type II pneumocytes, which may help treat or prevent disease such as hyaline membrane diseases, such as infant respiratory distress syndrome and bronchopulmonary dysplasia, in premature infants.

The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could stimulate the proliferation and differentiation of hepatocytes and, thus, could be used to alleviate or treat liver diseases and pathologies such as fulminant liver failure caused by cirrhosis, liver damage caused by viral hepatitis and

toxic substances (i.e., acetaminophen, carbon tetrachloride and other hepatotoxins known in the art).

In addition, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to treat or prevent the onset of diabetes mellitus. In patients with newly diagnosed Types I and II diabetes, where some islet cell function remains, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to maintain the islet function so as to alleviate, delay or prevent permanent manifestation of the disease. Also, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used as an auxiliary in islet cell transplantation to improve or promote islet cell function.

#### Neurological Diseases

Nervous system diseases, disorders, and/or conditions, which can be treated, prevented, and/or diagnosed with the compositions of the invention (e.g., polypeptides, polynucleotides, and/or agonists or antagonists), include, but are not limited to, nervous system injuries, and diseases, disorders, and/or conditions which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated, prevented, and/or diagnosed in a patient (including human and non-human mammalian patients) according to the invention, include but are not limited to, the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems: (1) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia; (2) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries; (3) malignant lesions, in which a portion of the nervous system is destroyed or injured by malignant tissue which is either a nervous system associated malignancy or a malignancy derived from non-nervous system tissue; (4) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis; (5) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative

process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis (ALS); (6) lesions associated with nutritional diseases, disorders, and/or conditions, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration; (7) neurological lesions associated with systemic diseases including, but not limited to, diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis; (8) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and (9) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including, but not limited to, multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

In a preferred embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to protect neural cells from the damaging effects of cerebral hypoxia. According to this embodiment, the compositions of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with cerebral hypoxia. In one aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with cerebral ischemia. In another aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with cerebral infarction. In another aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose or prevent neural cell injury associated with a stroke. In a further aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with a heart attack.

The compositions of the invention which are useful for treating or preventing a nervous system disorder may be selected by testing for biological activity in

promoting the survival or differentiation of neurons. For example, and not by way of limitation, compositions of the invention which elicit any of the following effects may be useful according to the invention: (1) increased survival time of neurons in culture; (2) increased sprouting of neurons in culture or in vivo; (3) increased production of a neuron-associated molecule in culture or in vivo, e.g., choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or (4) decreased symptoms of neuron dysfunction in vivo. Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may routinely be measured using a method set forth herein or otherwise known in the art, such as, for example, the method set forth in Arakawa et al. (*J. Neurosci.* 10:3507-3515 (1990)); increased sprouting of neurons may be detected by methods known in the art, such as, for example, the methods set forth in Pestronk et al. (*Exp. Neurol.* 70:65-82 (1980)) or Brown et al. (*Ann. Rev. Neurosci.* 4:17-42 (1981)); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., using techniques known in the art and depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron diseases, disorders, and/or conditions that may be treated, prevented, and/or diagnosed according to the invention include, but are not limited to, diseases, disorders, and/or conditions such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as diseases, disorders, and/or conditions that selectively affect neurons such as amyotrophic lateral sclerosis, and including, but not limited to, progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motor Sensory Neuropathy (Charcot-Marie-Tooth Disease).

#### Infectious Disease

A polypeptide or polynucleotide and/or agonist or antagonist of the present invention can be used to treat, prevent, and/or diagnose infectious agents. For

example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated, prevented, and/or diagnosed. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively,  
 5 polypeptide or polynucleotide and/or agonist or antagonist of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated, prevented, and/or diagnosed by a polynucleotide or  
 10 polypeptide and/or agonist or antagonist of the present invention. Examples of viruses, include, but are not limited to Examples of viruses, include, but are not limited to the following DNA and RNA viruses and viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Dengue, EBV, HIV, Flaviviridae, Hepadnaviridae  
 15 (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza A, Influenza B, and parainfluenza), Papiloma virus, Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus),  
 20 and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, respiratory syncytial virus, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), Japanese B encephalitis, Junin, Chikungunya, Rift Valley fever, yellow fever, meningitis,  
 25 opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. polynucleotides or polypeptides, or agonists or antagonists of the invention, can be used to treat, prevent, and/or diagnose any of these symptoms or  
 30 diseases. In specific embodiments, polynucleotides, polypeptides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose: meningitis, Dengue, EBV, and/or hepatitis (e.g., hepatitis B). In an additional specific

embodiment polynucleotides, polypeptides, or agonists or antagonists of the invention are used to treat patients nonresponsive to one or more other commercially available hepatitis vaccines. In a further specific embodiment polynucleotides, polypeptides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose  
 5 AIDS.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated, prevented, and/or diagnosed by a polynucleotide or polypeptide and/or agonist or antagonist of the present invention include, but not limited to, include, but not limited to, the following Gram-Negative and Gram-positive bacteria  
 10 and bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), Cryptococcus neoformans, Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia (e.g., Borrelia burgdorferi), Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, E. coli (e.g., Enterotoxigenic  
 15 E. coli and Enterohemorrhagic E. coli), Enterobacteriaceae (Klebsiella, Salmonella (e.g., Salmonella typhi, and Salmonella paratyphi), Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Mycobacterium leprae, Vibrio cholerae, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Meisseria meningitidis, Pasteurellacea Infections (e.g.,  
 20 Actinobacillus, Heamophilus (e.g., Heamophilus influenza type B), Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, Shigella spp., Staphylococcal, Meningiococcal, Pneumococcal and Streptococcal (e.g., Streptococcus pneumoniae and Group B Streptococcus). These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to:  
 25 bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis (e.g., meningitis types A and B), Chlamydia, Syphilis, Diphtheria, Leprosy,  
 30 Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g.,

cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. Polynucleotides or polypeptides, agonists or antagonists of the invention, can be used to treat, prevent, and/or diagnose any of these symptoms or diseases. In specific embodiments, polynucleotides, polypeptides, agonists or antagonists of the invention are used to treat, prevent, and/or diagnose: tetanus, Diphtheria, botulism, and/or meningitis type B.

Moreover, parasitic agents causing disease or symptoms that can be treated, prevented, and/or diagnosed by a polynucleotide or polypeptide and/or agonist or antagonist of the present invention include, but not limited to, the following families or class: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, 10 Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas and Sporozoans (e.g., Plasmodium virax, Plasmodium falciparum, Plasmodium malariae and Plasmodium ovale). These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, 15 giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), malaria, pregnancy complications, and toxoplasmosis. polynucleotides or polypeptides, or agonists or antagonists of the invention, can be used to treat, prevent, and/or diagnose any of these symptoms or diseases. In specific embodiments, polynucleotides, polypeptides, or agonists or antagonists of the invention are used to 20 treat, prevent, and/or diagnose malaria.

Preferably, treatment or prevention using a polypeptide or polynucleotide and/or agonist or antagonist of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and 25 returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

#### Regeneration

30 A polynucleotide or polypeptide and/or agonist or antagonist of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues

could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

5           Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

10           Moreover, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide and/or agonist or antagonist of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases  
15           that could be treated, prevented, and/or diagnosed include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

            Similarly, nerve and brain tissue could also be regenerated by using a  
20           polynucleotide or polypeptide and/or agonist or antagonist of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated, prevented, and/or diagnosed using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic diseases, disorders, and/or conditions (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and  
25           stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated, prevented, and/or diagnosed using the  
30           polynucleotide or polypeptide and/or agonist or antagonist of the present invention.

### Chemotaxis

A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat, prevent, and/or diagnose inflammation, infection, hyperproliferative diseases, disorders, and/or conditions, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat, prevent, and/or diagnose wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat, prevent, and/or diagnose wounds.

It is also contemplated that a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may inhibit chemotactic activity. These molecules could also be used to treat, prevent, and/or diagnose diseases, disorders, and/or conditions. Thus, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention could be used as an inhibitor of chemotaxis.

### Binding Activity

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural

receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

Preferably, the screening for these molecules involves producing appropriate  
5 cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of  
10 either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

15 Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a  
20 standard.

Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

25 Additionally, the receptor to which a polypeptide of the invention binds can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting (Coligan, et al., *Current Protocols in Immun.*, 1(2), Chapter 5, (1991)). For example, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the polypeptides, for  
30 example, NIH3T3 cells which are known to contain multiple receptors for the FGF family proteins, and SC-3 cells, and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the

polypeptides. Transfected cells which are grown on glass slides are exposed to the polypeptide of the present invention, after they have been labeled. The polypeptides can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase.

5           Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an iterative sub-pooling and re-screening process, eventually yielding a single clones that encodes the putative receptor.

          As an alternative approach for receptor identification, the labeled polypeptides  
10       can be photoaffinity linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE analysis and exposed to X-ray film. The labeled complex containing the receptors of the polypeptides can be excised, resolved into peptide fragments, and subjected to protein microsequencing. The amino acid sequence obtained from microsequencing would be  
15       used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the genes encoding the putative receptors.

          Moreover, the techniques of gene-shuffling, motif-shuffling, exon-shuffling, and/or codon-shuffling (collectively referred to as "DNA shuffling") may be employed to modulate the activities of polypeptides of the invention thereby  
20       effectively generating agonists and antagonists of polypeptides of the invention. See generally, U.S. Patent Nos. 5,605,793, 5,811,238, 5,830,721, 5,834,252, and 5,837,458, and Patten, P. A., et al., Curr. Opinion Biotechnol. 8:724-33 (1997); Harayama, S. Trends Biotechnol. 16(2):76-82 (1998); Hansson, L. O., et al., J. Mol. Biol. 287:265-76 (1999); and Lorenzo, M. M. and Blasco, R. Biotechniques  
25       24(2):308-13 (1998) (each of these patents and publications are hereby incorporated by reference). In one embodiment, alteration of polynucleotides and corresponding polypeptides of the invention may be achieved by DNA shuffling. DNA shuffling involves the assembly of two or more DNA segments into a desired polynucleotide sequence of the invention molecule by homologous, or site-specific, recombination.  
30       In another embodiment, polynucleotides and corresponding polypeptides of the invention may be altered by being subjected to random mutagenesis by error-prone PCR, random nucleotide insertion or other methods prior to recombination. In another

embodiment, one or more components, motifs, sections, parts, domains, fragments, etc., of the polypeptides of the invention may be recombined with one or more components, motifs, sections, parts, domains, fragments, etc. of one or more heterologous molecules. In preferred embodiments, the heterologous molecules are family members. In further preferred embodiments, the heterologous molecule is a growth factor such as, for example, platelet-derived growth factor (PDGF), insulin-like growth factor (IGF-I), transforming growth factor (TGF)-alpha, epidermal growth factor (EGF), fibroblast growth factor (FGF), TGF-beta, bone morphogenetic protein (BMP)-2, BMP-4, BMP-5, BMP-6, BMP-7, activins A and B, decapentaplegic(dpp), 60A, OP-2, dorsalin, growth differentiation factors (GDFs), nodal, MIS, inhibin-alpha, TGF-beta1, TGF-beta2, TGF-beta3, TGF-beta5, and glial-derived neurotrophic factor (GDNF).

Other preferred fragments are biologically active fragments of the polypeptides of the invention. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Additionally, this invention provides a method of screening compounds to identify those which modulate the action of the polypeptide of the present invention. An example of such an assay comprises combining a mammalian fibroblast cell, a the polypeptide of the present invention, the compound to be screened and 3[H] thymidine under cell culture conditions where the fibroblast cell would normally proliferate. A control assay may be performed in the absence of the compound to be screened and compared to the amount of fibroblast proliferation in the presence of the compound to determine if the compound stimulates proliferation by determining the uptake of 3[H] thymidine in each case. The amount of fibroblast cell proliferation is measured by liquid scintillation chromatography which measures the incorporation of 3[H] thymidine. Both agonist and antagonist compounds may be identified by this procedure.

In another method, a mammalian cell or membrane preparation expressing a receptor for a polypeptide of the present invention is incubated with a labeled polypeptide of the present invention in the presence of the compound. The ability of

the compound to enhance or block this interaction could then be measured. Alternatively, the response of a known second messenger system following interaction of a compound to be screened and the receptor is measured and the ability of the compound to bind to the receptor and elicit a second messenger response is measured to determine if the compound is a potential agonist or antagonist. Such second messenger systems include but are not limited to, cAMP guanylate cyclase, ion channels or phosphoinositide hydrolysis.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat, prevent, and/or diagnose disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptides of the invention from suitably manipulated cells or tissues. Therefore, the invention includes a method of identifying compounds which bind to the polypeptides of the invention comprising the steps of: (a) incubating a candidate binding compound with the polypeptide; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with the polypeptide, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

Also, one could identify molecules bind a polypeptide of the invention experimentally by using the beta-pleated sheet regions contained in the polypeptide sequence of the protein. Accordingly, specific embodiments of the invention are directed to polynucleotides encoding polypeptides which comprise, or alternatively consist of, the amino acid sequence of each beta pleated sheet regions in a disclosed polypeptide sequence. Additional embodiments of the invention are directed to polynucleotides encoding polypeptides which comprise, or alternatively consist of, any combination or all of contained in the polypeptide sequences of the invention. Additional preferred embodiments of the invention are directed to polypeptides which comprise, or alternatively consist of, the amino acid sequence of each of the beta pleated sheet regions in one of the polypeptide sequences of the invention. Additional embodiments of the invention are directed to polypeptides which comprise, or

alternatively consist of, any combination or all of the beta pleated sheet regions in one of the polypeptide sequences of the invention.

#### Targeted Delivery

5 In another embodiment, the invention provides a method of delivering compositions to targeted cells expressing a receptor for a polypeptide of the invention, or cells expressing a cell bound form of a polypeptide of the invention.

As discussed herein, polypeptides or antibodies of the invention may be associated with heterologous polypeptides, heterologous nucleic acids, toxins, or prodrugs via hydrophobic, hydrophilic, ionic and/or covalent interactions. In one 10 embodiment, the invention provides a method for the specific delivery of compositions of the invention to cells by administering polypeptides of the invention (including antibodies) that are associated with heterologous polypeptides or nucleic acids. In one example, the invention provides a method for delivering a therapeutic protein into the targeted cell. In another example, the invention provides a method for 15 delivering a single stranded nucleic acid (e.g., antisense or ribozymes) or double stranded nucleic acid (e.g., DNA that can integrate into the cell's genome or replicate episomally and that can be transcribed) into the targeted cell.

In another embodiment, the invention provides a method for the specific destruction of cells (e.g., the destruction of tumor cells) by administering polypeptides 20 of the invention (e.g., polypeptides of the invention or antibodies of the invention) in association with toxins or cytotoxic prodrugs.

By "toxin" is meant compounds that bind and activate endogenous cytotoxic effector systems, radioisotopes, holotoxins, modified toxins, catalytic subunits of toxins, or any molecules or enzymes not normally present in or on the surface of a cell 25 that under defined conditions cause the cell's death. Toxins that may be used according to the methods of the invention include, but are not limited to, radioisotopes known in the art, compounds such as, for example, antibodies (or complement fixing containing portions thereof) that bind an inherent or induced endogenous cytotoxic effector system, thymidine kinase, endonuclease, RNase, alpha toxin, ricin, abrin, 30 Pseudomonas exotoxin A, diphtheria toxin, saporin, momordin, gelonin, pokeweed antiviral protein, alpha-sarcin and cholera toxin. By "cytotoxic prodrug" is meant a non-toxic compound that is converted by an enzyme, normally present in the cell, into

a cytotoxic compound. Cytotoxic prodrugs that may be used according to the methods of the invention include, but are not limited to, glutamyl derivatives of benzoic acid mustard alkylating agent, phosphate derivatives of etoposide or mitomycin C, cytosine arabinoside, daunorubisin, and phenoxyacetamide derivatives of  
5 doxorubicin.

#### Drug Screening

Further contemplated is the use of the polypeptides of the present invention, or the polynucleotides encoding these polypeptides, to screen for molecules which modify the activities of the polypeptides of the present invention. Such a method  
10 would include contacting the polypeptide of the present invention with a selected compound(s) suspected of having antagonist or agonist activity, and assaying the activity of these polypeptides following binding.

This invention is particularly useful for screening therapeutic compounds by using the polypeptides of the present invention, or binding fragments thereof, in any  
15 of a variety of drug screening techniques. The polypeptide or fragment employed in such a test may be affixed to a solid support, expressed on a cell surface, free in solution, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or fragment. Drugs are screened against such transformed  
20 cells in competitive binding assays. One may measure, for example, the formulation of complexes between the agent being tested and a polypeptide of the present invention.

Thus, the present invention provides methods of screening for drugs or any other agents which affect activities mediated by the polypeptides of the present  
25 invention. These methods comprise contacting such an agent with a polypeptide of the present invention or a fragment thereof and assaying for the presence of a complex between the agent and the polypeptide or a fragment thereof, by methods well known in the art. In such a competitive binding assay, the agents to screen are typically labeled. Following incubation, free agent is separated from that present in bound  
30 form, and the amount of free or uncomplexed label is a measure of the ability of a particular agent to bind to the polypeptides of the present invention.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to the polypeptides of the present invention, and is described in great detail in European Patent Application 84/03564, published on September 13, 1984, which is incorporated herein by reference herein.

5 Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with polypeptides of the present invention and washed. Bound polypeptides are then detected by methods well known in the art. Purified polypeptides are coated directly onto plates for use in the aforementioned  
10 drug screening techniques. In addition, non-neutralizing antibodies may be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding polypeptides of the present invention specifically compete with a test compound for binding to the polypeptides  
15 or fragments thereof. In this manner, the antibodies are used to detect the presence of any peptide which shares one or more antigenic epitopes with a polypeptide of the invention.

The human phosphatase polypeptides and/or peptides of the present invention, or immunogenic fragments or oligopeptides thereof, can be used for screening  
20 therapeutic drugs or compounds in a variety of drug screening techniques. The fragment employed in such a screening assay may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The reduction or abolition of activity of the formation of binding complexes between the ion channel protein and the agent being tested can be measured. Thus, the present invention  
25 provides a method for screening or assessing a plurality of compounds for their specific binding affinity with a phosphatase polypeptide, or a bindable peptide fragment, of this invention, comprising providing a plurality of compounds, combining the phosphatase polypeptide, or a bindable peptide fragment, with each of a plurality of compounds for a time sufficient to allow binding under suitable  
30 conditions and detecting binding of the phosphatase polypeptide or peptide to each of the plurality of test compounds, thereby identifying the compounds that specifically bind to the phosphatase polypeptide or peptide.

Methods of identifying compounds that modulate the activity of the novel human phosphatase polypeptides and/or peptides are provided by the present invention and comprise combining a potential or candidate compound or drug modulator of phosphatase activity with a phosphatase polypeptide or peptide, for example, the phosphatase amino acid sequence as set forth in SEQ ID NO:42, 109, 150, or 152, and measuring an effect of the candidate compound or drug modulator on the biological activity of the phosphatase polypeptide or peptide. Such measurable effects include, for example, physical binding interaction; the ability to phosphorylate a suitable calpain substrate; effects on native and cloned phosphatase-expressing cell line; and effects of modulators or other phosphatase-mediated physiological measures.

Another method of identifying compounds that modulate the biological activity of the novel phosphatase polypeptides of the present invention comprises combining a potential or candidate compound or drug modulator of a phosphatase activity with a host cell that expresses the phosphatase polypeptide and measuring an effect of the candidate compound or drug modulator on the biological activity of the phosphatase polypeptide. The host cell can also be capable of being induced to express the phosphatase polypeptide, e.g., via inducible expression. Physiological effects of a given modulator candidate on the phosphatase polypeptide can also be measured. Thus, cellular assays for particular phosphatase modulators may be either direct measurement or quantification of the physical biological activity of the phosphatase polypeptide, or they may be measurement or quantification of a physiological effect. Such methods preferably employ a phosphatase polypeptide as described herein, or an overexpressed recombinant phosphatase polypeptide in suitable host cells containing an expression vector as described herein, wherein the phosphatase polypeptide is expressed, overexpressed, or undergoes upregulated expression.

Another aspect of the present invention embraces a method of screening for a compound that is capable of modulating the biological activity of a phosphatase polypeptide, comprising providing a host cell containing an expression vector harboring a nucleic acid sequence encoding a phosphatase polypeptide, or a functional peptide or portion thereof (e.g., SEQ ID NO:42, 109, 150, or 152); determining the biological activity of the expressed phosphatase polypeptide in the absence of a

modulator compound; contacting the cell with the modulator compound and determining the biological activity of the expressed phosphatase polypeptide in the presence of the modulator compound. In such a method, a difference between the activity of the phosphatase polypeptide in the presence of the modulator compound and in the absence of the modulator compound indicates a modulating effect of the compound.

Essentially any chemical compound can be employed as a potential modulator or ligand in the assays according to the present invention. Compounds tested as phosphatase modulators can be any small chemical compound, or biological entity (e.g., protein, sugar, nucleic acid, lipid). Test compounds will typically be small chemical molecules and peptides. Generally, the compounds used as potential modulators can be dissolved in aqueous or organic (e.g., DMSO-based) solutions. The assays are designed to screen large chemical libraries by automating the assay steps and providing compounds from any convenient source. Assays are typically run in parallel, for example, in microtiter formats on microtiter plates in robotic assays. There are many suppliers of chemical compounds, including Sigma (St. Louis, MO), Aldrich (St. Louis, MO), Sigma-Aldrich (St. Louis, MO), Fluka Chemika-Biochemica Analytika (Buchs, Switzerland), for example. Also, compounds may be synthesized by methods known in the art.

High throughput screening methodologies are particularly envisioned for the detection of modulators of the novel phosphatase polynucleotides and polypeptides described herein. Such high throughput screening methods typically involve providing a combinatorial chemical or peptide library containing a large number of potential therapeutic compounds (e.g., ligand or modulator compounds). Such combinatorial chemical libraries or ligand libraries are then screened in one or more assays to identify those library members (e.g., particular chemical species or subclasses) that display a desired characteristic activity. The compounds so identified can serve as conventional lead compounds, or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated either by chemical synthesis or biological synthesis, by combining a number of chemical building blocks (i.e., reagents such as amino acids).

As an example, a linear combinatorial library, e.g., a polypeptide or peptide library, is formed by combining a set of chemical building blocks in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide or peptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks.

The preparation and screening of combinatorial chemical libraries is well known to those having skill in the pertinent art. Combinatorial libraries include, without limitation, peptide libraries (e.g. U.S. Patent No. 5,010,175; Furka, 1991, *Int. J. Pept. Prot. Res.*, 37:487-493; and Houghton et al., 1991, *Nature*, 354:84-88). Other chemistries for generating chemical diversity libraries can also be used. Nonlimiting examples of chemical diversity library chemistries include, peptides (PCT Publication No. WO 91/019735), encoded peptides (PCT Publication No. WO 93/20242), random bio-oligomers (PCT Publication No. WO 92/00091), benzodiazepines (U.S. Patent No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs et al., 1993, *Proc. Natl. Acad. Sci. USA*, 90:6909-6913), vinylogous polypeptides (Hagihara et al., 1992, *J. Amer. Chem. Soc.*, 114:6568), nonpeptidal peptidomimetics with glucose scaffolding (Hirschmann et al., 1992, *J. Amer. Chem. Soc.*, 114:9217-9218), analogous organic synthesis of small compound libraries (Chen et al., 1994, *J. Amer. Chem. Soc.*, 116:2661), oligocarbamates (Cho et al., 1993, *Science*, 261:1303), and/or peptidyl phosphonates (Campbell et al., 1994, *J. Org. Chem.*, 59:658), nucleic acid libraries (see Ausubel, Berger and Sambrook, all supra), peptide nucleic acid libraries (U.S. Patent No. 5,539,083), antibody libraries (e.g., Vaughn et al., 1996, *Nature Biotechnology*, 14(3):309-314) and PCT/US96/10287), carbohydrate libraries (e.g., Liang et al., 1996, *Science*, 274:1520-1522) and U.S. Patent No. 5,593,853), small organic molecule libraries (e.g., benzodiazepines, Baum C&EN, Jan. 18, 1993, page 33; and U.S. Patent No. 5,288,514; isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; and the like).

Devices for the preparation of combinatorial libraries are commercially available (e.g., 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY; Symphony, Rainin, Woburn, MA; 433A Applied Biosystems, Foster City, CA; 9050

Plus, Millipore, Bedford, MA). In addition, a large number of combinatorial libraries are commercially available (e.g., ComGenex, Princeton, NJ; Asinex, Moscow, Russia; Tripos, Inc., St. Louis, MO; ChemStar, Ltd., Moscow, Russia; 3D Pharmaceuticals, Exton, PA; Martek Biosciences, Columbia, MD, and the like).

5           In one embodiment, the invention provides solid phase based *in vitro* assays in a high throughput format, where the cell or tissue expressing an ion channel is attached to a solid phase substrate. In such high throughput assays, it is possible to screen up to several thousand different modulators or ligands in a single day. In particular, each well of a microtiter plate can be used to perform a separate assay  
10       against a selected potential modulator, or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 96 modulators. If 1536 well plates are used, then a single plate can easily assay from about 100 to about 1500 different compounds. It is possible to assay several different plates per day; thus, for example,  
15       assay screens for up to about 6,000-20,000 different compounds are possible using the described integrated systems.

          In another of its aspects, the present invention encompasses screening and small molecule (e.g., drug) detection assays which involve the detection or identification of small molecules that can bind to a given protein, i.e., a phosphatase  
20       polypeptide or peptide. Particularly preferred are assays suitable for high throughput screening methodologies.

          In such binding-based detection, identification, or screening assays, a functional assay is not typically required. All that is needed is a target protein, preferably substantially purified, and a library or panel of compounds (e.g., ligands,  
25       drugs, small molecules) or biological entities to be screened or assayed for binding to the protein target. Preferably, most small molecules that bind to the target protein will modulate activity in some manner, due to preferential, higher affinity binding to functional areas or sites on the protein.

          An example of such an assay is the fluorescence based thermal shift assay (3-  
30       Dimensional Pharmaceuticals, Inc., 3DP, Exton, PA) as described in U.S. Patent Nos. 6,020,141 and 6,036,920 to Pantoliano et al.; see also, J. Zimmerman, 2000, *Gen. Eng. News*, 20(8)). The assay allows the detection of small molecules (e.g., drugs,

ligands) that bind to expressed, and preferably purified, ion channel polypeptide based on affinity of binding determinations by analyzing thermal unfolding curves of protein-drug or ligand complexes. The drugs or binding molecules determined by this technique can be further assayed, if desired, by methods, such as those described  
5 herein, to determine if the molecules affect or modulate function or activity of the target protein.

To purify a phosphatase polypeptide or peptide to measure a biological binding or ligand binding activity, the source may be a whole cell lysate that can be prepared by successive freeze-thaw cycles (e.g., one to three) in the presence of  
10 standard protease inhibitors. The phosphatase polypeptide may be partially or completely purified by standard protein purification methods, e.g., affinity chromatography using specific antibody described *infra*, or by ligands specific for an epitope tag engineered into the recombinant phosphatase polypeptide molecule, also as described herein. Binding activity can then be measured as described.

15 Compounds which are identified according to the methods provided herein, and which modulate or regulate the biological activity or physiology of the phosphatase polypeptides according to the present invention are a preferred embodiment of this invention. It is contemplated that such modulatory compounds may be employed in treatment and therapeutic methods for treating a condition that is  
20 mediated by the novel phosphatase polypeptides by administering to an individual in need of such treatment a therapeutically effective amount of the compound identified by the methods described herein.

In addition, the present invention provides methods for treating an individual in need of such treatment for a disease, disorder, or condition that is mediated by the  
25 phosphatase polypeptides of the invention, comprising administering to the individual a therapeutically effective amount of the phosphatase-modulating compound identified by a method provided herein.

#### Antisense And Ribozyme (Antagonists)

In specific embodiments, antagonists according to the present invention are  
30 nucleic acids corresponding to the sequences contained in SEQ ID NO:1, or the complementary strand thereof, and/or to nucleotide sequences contained a deposited clone. In one embodiment, antisense sequence is generated internally by the organism,

in another embodiment, the antisense sequence is separately administered (see, for example, O'Connor, *Neurochem.*, 56:560 (1991). *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*, CRC Press, Boca Raton, FL (1988). Antisense technology can be used to control gene expression through antisense DNA or RNA, or through triple-helix formation. Antisense techniques are discussed for example, in Okano, *Neurochem.*, 56:560 (1991); *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*, CRC Press, Boca Raton, FL (1988). Triple helix formation is discussed in, for instance, Lee et al., *Nucleic Acids Research*, 6:3073 (1979); Cooney et al., *Science*, 241:456 (1988); and Dervan et al., *Science*, 251:1300 (1991). The methods are based on binding of a polynucleotide to a complementary DNA or RNA.

For example, the use of c-myc and c-myb antisense RNA constructs to inhibit the growth of the non-lymphocytic leukemia cell line HL-60 and other cell lines was previously described. (Wickstrom et al. (1988); Anfossi et al. (1989)). These experiments were performed in vitro by incubating cells with the oligoribonucleotide. A similar procedure for in vivo use is described in WO 91/15580. Briefly, a pair of oligonucleotides for a given antisense RNA is produced as follows: A sequence complimentary to the first 15 bases of the open reading frame is flanked by an EcoRI site on the 5' end and a HindIII site on the 3' end. Next, the pair of oligonucleotides is heated at 90°C for one minute and then annealed in 2X ligation buffer (20mM TRIS HCl pH 7.5, 10mM MgCl<sub>2</sub>, 10mM dithiothreitol (DTT) and 0.2 mM ATP) and then ligated to the EcoRI/Hind III site of the retroviral vector PMV7 (WO 91/15580).

For example, the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of the receptor. The antisense RNA oligonucleotide hybridizes to the mRNA in vivo and blocks translation of the mRNA molecule into receptor polypeptide. Antisense oligonucleotides may be single or double stranded. Double stranded RNA's may be designed based upon the teachings of Paddison et al., *Proc. Nat. Acad. Sci.*, 99:1443-1448 (2002); and International Publication Nos. WO 01/29058, and WO 99/32619; which are hereby incorporated herein by reference.

In one embodiment, the antisense nucleic acid of the invention is produced intracellularly by transcription from an exogenous sequence. For example, a vector or a portion thereof, is transcribed, producing an antisense nucleic acid (RNA) of the invention. Such a vector would contain a sequence encoding the antisense nucleic acid of the invention. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in vertebrate cells. Expression of the sequence encoding a polypeptide of the invention, or fragments thereof, can be by any promoter known in the art to act in vertebrate, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include, but are not limited to, the SV40 early promoter region (Bernoist and Chambon, *Nature*, 29:304-310 (1981)), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., *Cell*, 22:787-797 (1980)), the herpes thymidine promoter (Wagner et al., *Proc. Natl. Acad. Sci. U.S.A.*, 78:1441-1445 (1981)), the regulatory sequences of the metallothionein gene (Brinster et al., *Nature*, 296:39-42 (1982)), etc.

The antisense nucleic acids of the invention comprise a sequence complementary to at least a portion of an RNA transcript of a gene of interest. However, absolute complementarity, although preferred, is not required. A sequence "complementary to at least a portion of an RNA," referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double stranded antisense nucleic acids of the invention, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the larger the hybridizing nucleic acid, the more base mismatches with a RNA sequence of the invention it may contain and still form a stable duplex (or triplex as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the message, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should

work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have been shown to be effective at inhibiting translation of mRNAs as well. See generally, Wagner, R., *Nature*, 372:333-335 (1994). Thus, oligonucleotides complementary to either the 5' - or 3' - non-translated, non-coding regions of a polynucleotide sequence of the invention could be used in an antisense approach to inhibit translation of endogenous mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could be used in accordance with the invention. Whether designed to hybridize to the 5' -, 3' - or coding region of mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides or at least 50 nucleotides.

The polynucleotides of the invention can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556 (1989); Lemaitre et al., *Proc. Natl. Acad. Sci.*, 84:648-652 (1987); PCT Publication NO: WO88/09810, published December 15, 1988) or the blood-brain barrier (see, e.g., PCT Publication NO: WO89/10134, published April 25, 1988), hybridization-triggered cleavage agents. (See, e.g., Krol et al., *BioTechniques*, 6:958-976 (1988)) or intercalating agents. (See, e.g., Zon, *Pharm. Res.*, 5:539-549 (1988)). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine,

5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including, but not limited to, arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group including, but not limited to, a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an a-anomeric oligonucleotide. An a-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual b-units, the strands run parallel to each other (Gautier et al., Nucl. Acids Res., 15:6625-6641 (1987)). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., Nucl. Acids Res., 15:6131-6148 (1987)), or a chimeric RNA-DNA analogue (Inoue et al., FEBS Lett. 215:327-330 (1987)).

Polynucleotides of the invention may be synthesized by standard methods known in the art, e.g. by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein et al. (Nucl. Acids Res., 16:3209 (1988)), methylphosphonate oligonucleotides can be

prepared by use of controlled pore glass polymer supports (Sarin et al., Proc. Natl. Acad. Sci. U.S.A., 85:7448-7451 (1988)), etc.

While antisense nucleotides complementary to the coding region sequence of the invention could be used, those complementary to the transcribed untranslated  
5 region are most preferred.

Potential antagonists according to the invention also include catalytic RNA, or a ribozyme (See, e.g., PCT International Publication WO 90/11364, published October 4, 1990; Sarver et al, Science, 247:1222-1225 (1990). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy mRNAs  
10 corresponding to the polynucleotides of the invention, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5' -UG-3'. The construction and production of hammerhead ribozymes is well known  
15 in the art and is described more fully in Haseloff and Gerlach, Nature, 334:585-591 (1988). There are numerous potential hammerhead ribozyme cleavage sites within each nucleotide sequence disclosed in the sequence listing. Preferably, the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the mRNA corresponding to the polynucleotides of the invention; i.e., to increase  
20 efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

As in the antisense approach, the ribozymes of the invention can be composed of modified oligonucleotides (e.g. for improved stability, targeting, etc.) and should be delivered to cells which express the polynucleotides of the invention in vivo. DNA  
25 constructs encoding the ribozyme may be introduced into the cell in the same manner as described above for the introduction of antisense encoding DNA. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive promoter, such as, for example, pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to  
30 destroy endogenous messages and inhibit translation. Since ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Antagonist/agonist compounds may be employed to inhibit the cell growth and proliferation effects of the polypeptides of the present invention on neoplastic cells and tissues, i.e. stimulation of angiogenesis of tumors, and, therefore, retard or prevent abnormal cellular growth and proliferation, for example, in tumor formation or growth.

The antagonist/agonist may also be employed to prevent hyper-vascular diseases, and prevent the proliferation of epithelial lens cells after extracapsular cataract surgery. Prevention of the mitogenic activity of the polypeptides of the present invention may also be desirable in cases such as restenosis after balloon angioplasty.

The antagonist/agonist may also be employed to prevent the growth of scar tissue during wound healing.

The antagonist/agonist may also be employed to treat, prevent, and/or diagnose the diseases described herein.

Thus, the invention provides a method of treating or preventing diseases, disorders, and/or conditions, including but not limited to the diseases, disorders, and/or conditions listed throughout this application, associated with overexpression of a polynucleotide of the present invention by administering to a patient (a) an antisense molecule directed to the polynucleotide of the present invention, and/or (b) a ribozyme directed to the polynucleotide of the present invention.

#### Biotic Associations

A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase the organisms ability, either directly or indirectly, to initiate and/or maintain biotic associations with other organisms. Such associations may be symbiotic, nonsymbiotic, endosymbiotic, macrosymbiotic, and/or microsymbiotic in nature. In general, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase the organisms ability to form biotic associations with any member of the fungal, bacterial, lichen, mycorrhizal, cyanobacterial, dinoflagellate, and/or algal, kingdom, phylums, families, classes, genuses, and/or species.

The mechanism by which a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase the host organisms ability, either

directly or indirectly, to initiate and/or maintain biotic associations is variable, though may include, modulating osmolarity to desirable levels for the symbiont, modulating pH to desirable levels for the symbiont, modulating secretions of organic acids, modulating the secretion of specific proteins, phenolic compounds, nutrients, or the  
5 increased expression of a protein required for host-biotic organisms interactions (e.g., a receptor, ligand, etc.). Additional mechanisms are known in the art and are encompassed by the invention (see, for example, "Microbial Signalling and Communication", eds., R. England, G. Hobbs, N. Bainton, and D. McL. Roberts, Cambridge University Press, Cambridge, (1999); which is hereby incorporated herein  
10 by reference).

In an alternative embodiment, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may decrease the host organisms ability to form biotic associations with another organism, either directly or indirectly. The mechanism by which a polynucleotide or polypeptide and/or agonist or antagonist of  
15 the present invention may decrease the host organisms ability, either directly or indirectly, to initiate and/or maintain biotic associations with another organism is variable, though may include, modulating osmolarity to undesirable levels, modulating pH to undesirable levels, modulating secretions of organic acids, modulating the secretion of specific proteins, phenolic compounds, nutrients, or the  
20 decreased expression of a protein required for host-biotic organisms interactions (e.g., a receptor, ligand, etc.). Additional mechanisms are known in the art and are encompassed by the invention (see, for example, "Microbial Signalling and Communication", eds., R. England, G. Hobbs, N. Bainton, and D. McL. Roberts, Cambridge University Press, Cambridge, (1999); which is hereby incorporated herein  
25 by reference).

The hosts ability to maintain biotic associations with a particular pathogen has significant implications for the overall health and fitness of the host. For example, human hosts have symbiosis with enteric bacteria in their gastrointestinal tracts, particularly in the small and large intestine. In fact, bacteria counts in feces of the  
30 distal colon often approach  $10^{12}$  per milliliter of feces. Examples of bowel flora in the gastrointestinal tract are members of the Enterobacteriaceae, Bacteriodes, in addition to a-hemolytic streptococci, E. coli, Bifobacteria, Anaerobic cocci, Eubacteria,

Costridia, lactobacilli, and yeasts. Such bacteria, among other things, assist the host in the assimilation of nutrients by breaking down food stuffs not typically broken down by the hosts digestive system, particularly in the hosts bowel. Therefore, increasing the hosts ability to maintain such a biotic association would help assure proper  
5 nutrition for the host.

Aberrations in the enteric bacterial population of mammals, particularly humans, has been associated with the following disorders: diarrhea, ileus, chronic inflammatory disease, bowel obstruction, duodenal diverticula, biliary calculous disease, and malnutrition. A polynucleotide or polypeptide and/or agonist or  
10 antagonist of the present invention are useful for treating, detecting, diagnosing, prognosing, and/or ameliorating, either directly or indirectly, and of the above mentioned diseases and/or disorders associated with aberrant enteric flora population.

The composition of the intestinal flora, for example, is based upon a variety of factors, which include, but are not limited to, the age, race, diet, malnutrition, gastric  
15 acidity, bile salt excretion, gut motility, and immune mechanisms. As a result, the polynucleotides and polypeptides, including agonists, antagonists, and fragments thereof, may modulate the ability of a host to form biotic associations by affecting, directly or indirectly, at least one or more of these factors.

Although the predominate intestinal flora comprises anaerobic organisms, an  
20 underlying percentage represents aerobes (e.g., *E. coli*). This is significant as such aerobes rapidly become the predominate organisms in intraabdominal infections – effectively becoming opportunistic early in infection pathogenesis. As a result, there is an intrinsic need to control aerobe populations, particularly for immune compromised individuals.

25 In a preferred embodiment, a polynucleotides and polypeptides, including agonists, antagonists, and fragments thereof, are useful for inhibiting biotic associations with specific enteric symbiont organisms in an effort to control the population of such organisms.

Biotic associations occur not only in the gastrointestinal tract, but also on an in  
30 the integument. As opposed to the gastrointestinal flora, the cutaneous flora is comprised almost equally with aerobic and anaerobic organisms. Examples of cutaneous flora are members of the gram-positive cocci (e.g., *S. aureus*, coagulase-

negative staphylococci, micrococcus, *M.sedentarius*), gram-positive bacilli (e.g., *Corynebacterium* species, *C. minutissimum*, *Brevibacterium* species, *Propionibacterium* species, *P.acnes*), gram-negative bacilli (e.g., *Acinebacter* species), and fungi (*Pityrosporum orbiculare*). The relatively low number of flora associated with the integument is based upon the inability of many organisms to adhere to the skin. The organisms referenced above have acquired this unique ability. Therefore, the polynucleotides and polypeptides of the present invention may have uses which include modulating the population of the cutaneous flora, either directly or indirectly.

Aberrations in the cutaneous flora are associated with a number of significant diseases and/or disorders, which include, but are not limited to the following: impetigo, ecthyma, blistering distal dactulitis, pustules, folliculitis, cutaneous abscesses, pitted keratolysis, trichomycosis axcillaris, dermatophytosis complex, axillary odor, erthyrasma, cheesy foot odor, acne, tinea versicolor, seborrheic dermatitis, and *Pityrosporum* folliculitis, to name a few. A polynucleotide or polypeptide and/or agonist or antagonist of the present invention are useful for treating, detecting, diagnosing, prognosing, and/or ameliorating, either directly or indirectly, and of the above mentioned diseases and/or disorders associated with aberrant cutaneous flora population.

Additional biotic associations, including diseases and disorders associated with the aberrant growth of such associations, are known in the art and are encompassed by the invention. See, for example, "Infectious Disease", Second Edition, Eds., S.L., Gorbach, J.G., Bartlett, and N.R., Blacklow, W.B. Saunders Company, Philadelphia, (1998); which is hereby incorporated herein by reference).

Pheromones

In another embodiment, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase the organisms ability to synthesize and/or release a pheromone. Such a pheromone may, for example, alter the organisms behavior and/or metabolism.

A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may modulate the biosynthesis and/or release of pheromones, the organisms ability to respond to pheromones (e.g., behaviorally, and/or metabolically), and/or the

organisms ability to detect pheromones. Preferably, any of the pheromones, and/or volatiles released from the organism, or induced, by a polynucleotide or polypeptide and/or agonist or antagonist of the invention have behavioral effects the organism.

#### Other Activities

5           The polypeptide of the present invention, as a result of the ability to stimulate vascular endothelial cell growth, may be employed in treatment for stimulating re-vascularization of ischemic tissues due to various disease conditions such as thrombosis, arteriosclerosis, and other cardiovascular conditions. These polypeptide may also be employed to stimulate angiogenesis and limb regeneration, as discussed  
10   above.

          The polypeptide may also be employed for treating wounds due to injuries, burns, post-operative tissue repair, and ulcers since they are mitogenic to various cells of different origins, such as fibroblast cells and skeletal muscle cells, and therefore, facilitate the repair or replacement of damaged or diseased tissue.

15           The polypeptide of the present invention may also be employed stimulate neuronal growth and to treat, prevent, and/or diagnose neuronal damage which occurs in certain neuronal disorders or neuro-degenerative conditions such as Alzheimer's disease, Parkinson's disease, and AIDS-related complex. The polypeptide of the invention may have the ability to stimulate chondrocyte growth, therefore, they may  
20   be employed to enhance bone and periodontal regeneration and aid in tissue transplants or bone grafts.

          The polypeptide of the present invention may be also be employed to prevent skin aging due to sunburn by stimulating keratinocyte growth.

          The polypeptide of the invention may also be employed for preventing hair  
25   loss, since FGF family members activate hair-forming cells and promotes melanocyte growth. Along the same lines, the polypeptides of the present invention may be employed to stimulate growth and differentiation of hematopoietic cells and bone marrow cells when used in combination with other cytokines.

          The polypeptide of the invention may also be employed to maintain organs  
30   before transplantation or for supporting cell culture of primary tissues.

          The polypeptide of the present invention may also be employed for inducing tissue of mesodermal origin to differentiate in early embryos.

The polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

5 The polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, polypeptides or polynucleotides and/or agonist or antagonists of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, 10 utilization, and storage of energy.

Polypeptide or polynucleotides and/or agonist or antagonists of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, cardiac rhythms, depression (including depressive diseases, disorders, and/or conditions), tendency for violence, tolerance for pain, reproductive 15 capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

Polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, 20 minerals, cofactors or other nutritional components.

Polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used to increase the efficacy of a pharmaceutical composition, either directly or indirectly. Such a use may be administered in simultaneous conjunction with said pharmaceutical, or separately through either the same or 25 different route of administration (e.g., intravenous for the polynucleotide or polypeptide of the present invention, and orally for the pharmaceutical, among others described herein.).

Polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used to prepare individuals for extraterrestrial travel, low 30 gravity environments, prolonged exposure to extraterrestrial radiation levels, low oxygen levels, reduction of metabolic activity, exposure to extraterrestrial pathogens, etc. Such a use may be administered either prior to an extraterrestrial event, during an

extraterrestrial event, or both. Moreover, such a use may result in a number of beneficial changes in the recipient, such as, for example, any one of the following, non-limiting, effects: an increased level of hematopoietic cells, particularly red blood cells which would aid the recipient in coping with low oxygen levels; an increased  
5 level of B-cells, T-cells, antigen presenting cells, and/or macrophages, which would aid the recipient in coping with exposure to extraterrestrial pathogens, for example; a temporary (i.e., reversible) inhibition of hematopoietic cell production which would aid the recipient in coping with exposure to extraterrestrial radiation levels; increase and/or stability of bone mass which would aid the recipient in coping with low gravity  
10 environments; and/or decreased metabolism which would effectively facilitate the recipients ability to prolong their extraterrestrial travel by any one of the following, non-limiting means: (i) aid the recipient by decreasing their basal daily energy requirements; (ii) effectively lower the level of oxidative and/or metabolic stress in recipient (i.e., to enable recipient to cope with increased extraterrestrial radiation levels  
15 by decreasing the level of internal oxidative/metabolic damage acquired during normal basal energy requirements; and/or (iii) enabling recipient to subsist at a lower metabolic temperature (i.e., cryogenic, and/or sub-cryogenic environment).

#### Other Preferred Embodiments

Other preferred embodiments of the claimed invention include an isolated  
20 nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:1 wherein X is any integer as defined in Table I.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:1 in the range of  
25 positions beginning with the nucleotide at about the position of the “5′ NT of Start Codon of ORF” and ending with the nucleotide at about the position of the “3′ NT of ORF” as defined for SEQ ID NO:1 in Table I.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150  
30 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:1.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:1.

5 A further preferred embodiment is a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:1 beginning with the nucleotide at about the position of the "5' NT of ORF" and ending with the nucleotide at about the position of the "3' NT of ORF" as defined for SEQ ID NO:1 in Table I.

10 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:1.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a cDNA clone identified by a cDNA Clone Identifier in Table I, which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table I for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in the nucleotide sequence of a cDNA clone identified by a cDNA Clone Identifier in Table I, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table I.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said cDNA clone.

30 Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said cDNA clone.

5 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least  
10 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:1 wherein X is any integer as defined in Table I; and a nucleotide sequence encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I;  
15 which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences  
20 comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected  
25 from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least  
30 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:1 wherein X is any integer as defined in Table I; and a nucleotide sequence encoded by

a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a protein identified in Table I, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:1 wherein X is any integer as defined in Table I; and a nucleotide sequence encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:1 wherein X is any integer as defined in Table I; and a nucleotide sequence encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:2 wherein Y is any integer as defined in Table I.

Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:2 in the range of positions "Total AA of the Open Reading Frame (ORF)" as set forth for SEQ ID NO:2 in Table I.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:2.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:2.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:2.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the complete amino acid sequence of a protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of the protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of the protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in

the amino acid sequence of the protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

Also preferred is an isolated polypeptide comprising an amino acid sequence  
5 at least 95% identical to the amino acid sequence of the protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a  
10 sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:2 wherein Y is any integer as defined in Table I; and a complete amino acid sequence of a protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:2 wherein Y is any integer as defined in Table I; and a complete amino acid sequence of a protein encoded by a  
15 cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is at  
20 least 90% identical to said sequence of at least 10 contiguous amino acids.  
25

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide  
30 comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:2 wherein Y is any integer as defined in Table

I; and a complete amino acid sequence of a protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

Also preferred is the above method wherein said step of comparing sequences  
5 is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90%  
10 identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:2 wherein Y is any integer as defined in Table I; and a complete amino acid sequence of a protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone  
15 in Table I.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a  
20 sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing a pathological condition associated with an organism with abnormal structure or expression of a gene encoding a protein identified in Table I, which method comprises a step of detecting in a biological  
25 sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:2 wherein Y is any integer as defined in Table I; and a  
30 complete amino acid sequence of a protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:2 wherein Y is any integer as defined in Table I; and a complete amino acid sequence of a protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:2 wherein Y is any integer as defined in Table I; and a complete amino acid sequence of a protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I.

Further preferred is a method of making a recombinant vector comprising inserting any of the above isolated nucleic acid molecule(s) into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

Also preferred is a method of making an isolated polypeptide comprising culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:2 wherein Y is an integer set forth in Table I and said position of the "Total AA of ORF" of SEQ ID NO:2 is

defined in Table I; and an amino acid sequence of a protein encoded by a cDNA clone identified by a cDNA Clone Identifier in Table I and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table I. The isolated polypeptide produced by this method is also preferred.

5        Also preferred is a method of treatment of an individual in need of an increased level of a protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

10       Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

#### REFERENCES

15       Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., (more) 2000. The genome sequence of *Drosophila melanogaster*. *Science* 287:2185-2195.

Altschul, SF, Gish, W, Miller, W, Myers, EW, Lipman, DJ 1990. Basic local alignment search tool. *J Mol. Biol.* 215:403-410.

20       Altschul, SF, Madden, TL, Schaeffer, AA, Zhang, J, Zhang, Z, Miller, W, Lipman, DJ 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acid Res.* 25:3389-3402.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E. The Protein Data Bank *Nucleic Acids Research*,  
25       28:235-242, 2000

Bernstein, FC, Koetzle, TF, Williams, GJB, Meyer, EF Jr., Brice, MD, Rodgers, JR, Kennard, O, Simanouchi, T, Tasumi, M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535-542.

30       Bohm, H-J., LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comp. Aid. Molec. Design* 6:61-78, 1992.

- Cardozo, T., Totrov, M., Abagyan, R. Homology modeling by the ICM method. *Proteins* **23**:403-14, 1995.
- Drews, J. *Nature Biotechnol.* 17:406 (1999).
- Drews, J. *Science* 287:1960 (2000).
- 5 Fischer, D. & Eisenberg, D. 1996. Protein fold recognition using sequence derived predictions. *Protein Sci.* 5:947955.
- Fischer, D. & Eisenberg D. 1997. Assigning folds to the proteins encoded by the genome of *Mycoplasma genitalium*. *Proc. Natl. Acad. Sci., USA* 94:11929-11934.
- Goodford, P.J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* 10 **28**:849-857, 1985.
- Goodsell, D.S. and Olsen, A.J. Automated docking of substrates to proteins by simulated annealing. *Proteins* **8**:195-202, 1990.
- Greer, J. Comparative modeling of homologous proteins. *Meth.* 15 *Enzymol.* **202**:239-52, 1991.
- Godzik, A., Kolinski, A., & Skolnick, J. 1992. Topology fingerprint approach to the inverse folding problem *J. Mol. Biol.* 227:227-238.
- Hendlich, M., Lackner, P., Weitkus, S., Floeckner, H., Froschauer, R., Gottsbacher, K., Casari, G., Sippl, M. J. Identification of native protein folds amongst a large number of incorrect models. The calculation of low energy conformations 20 from potentials of mean force. *J. Mol. Biol.* **216**:167-80, 1990.
- Koppensteiner, W.A., lackner, P., Wiederstein, M. & Sippl, M.J. 2000. Title J. *Mol. Biol.* 296:1139-1152.
- Kuntz, I. D., Blaney, J.M., Oatley, S.J., Langridge, R., and Ferrin, T.E. A 25 geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **161**:269-288, 1982.
- Lackner, P., Koppensteiner, W., Sippl, M.J., and Domingues, F.S. 2000. ProSup: a refined tool for protein structure alignment. *Protein Engineering* 13:745-752.
- 30 Lesk, A. M., Boswell, D. R. Homology Modeling: Inferences from Tables of Aligned Sequences. *Curr. Op. Struc. Biol.* **2**: 242-247, 1992.

- Levitt, M. Accurate modeling of protein conformation by automatic segment matching *J. Mol. Biol.* **226**: 507-533, 1992.
- Luthy et al. *Nature* 356:83, 1992
- Jones, D.T., Taylor, W.R., & Thornton, J.M. 1992. A new approach to fold  
5 recognition. *Nature* 358:86-89.
- Martin, Y.C. 3D database searching in drug design. *J. Med. Chem.* **35**:2145-2154, 1992.
- Murzin, A.G., Brenner, S.E., Hubbard, T. & Chothia, C. 1995. SCOP: a  
structural classification of protein databases for the investigation of sequences and  
10 structures. *J. Mol. Biol.* 247:536-540.
- Novotny, J., Rashin, A. A., and Bruccoleri, R. E. Criteria that discriminate between native proteins and incorrectly folded models. *Proteins* **4**:19-30, 1988.
- Pearson, W.R. & Miller, W. 1992. Dynamic programming algorithms for biological sequence comparison. *Methods Enzymol.* 210:575-601.
- 15 Puius, Y.A., Zhao, Y., Sullivan, M., Lawrence, D.S., Almo, S.C., & Zhang, Z-Y. Identification of a second aryl phosphate-binding site in protein-tyrosine phosphatase 1B: A paradigm for inhibitor design. *Proc. Natl. Acad. Sci., USA* **94**:13420-13425.
- Rychlewski, L., Zhang, B.H., & Godzik, A. 1998 Fold and function  
20 predictions for *Mycoplasma genitalium* proteins. *Folding & Design* 3:229-238.
- Rychlewski, L., Zhang, B.H., & Godzik, A 1999. Functional insights from structural predictions: analysis of the *Escherichia coli* genome. *Protein Sci.* 8:614-624.
- Sali A; Potterton L; Yuan F; van Vlijmen H; Karplus M Evaluation of  
25 comparative protein modeling by MODELLER. *Proteins* 23:318-26 (1995).
- Sippl M.J. 1990. Calculation of conformational ensembles from potentials of mean force: an approach to knowledge-based prediction of local structures in globular proteins. *J. Mol. Biol.* 213:859-883.
- Sippl, M.J. & Weickus, S. 1992. Detection of Native-like models for amino  
30 acid sequences of unknown three-dimensional structure in a data base of known protein conformations. *Proteins:Structure, Function and Genetics* 13:258-271.

Sippl M.J. 1993. Boltzmann's principle, knowledge-based mean fields and protein folding. An approach to the computational determination of protein structures. *Comp. Aided Mol. Design* 7:473.

Xu et al. *Nature Structural Biol.* 6:750 1999.

- 5        Zu-Kang, F. & Sippl, M.J. 1996 Optimum superposition of protein structures: ambiguities and implications. *Folding & Design* 1:123-132.

Lander et al *Nature* 409:860, 2001.

## EXAMPLES

## Description of the Preferred Embodiments

EXAMPLE 1 – METHOD OF IDENTIFYING THE NOVEL BMY\_HPP HUMAN  
5 PHOSPHATASES OF THE PRESENT INVENTION

Polynucleotide sequences encoding the novel BMY\_HPP13 phosphatase of the present invention was identified by a combination of the structural threading and genomic mining methodologies.

10 Typical genome mining methods capitalize on sequence similarity as the basis for assigning gene function. However, the primary structure of a distantly related unknown gene (< 30 percent identity at the level of the amino acid sequence) cannot always yield predicatable information leading to the structure of that gene, and is not always a good predictor of function. However, if there is a correlation between amino acid sequence and protein structure the characterization of function becomes easier.

15 It has been shown that clear sequence similarity implies structural similarity (sequence identity >50%). In addition it has also been shown that structural similarity exists for proteins even when the sequence similarity lies in the “twilight zone” (<30% sequence identity). There are several methods for detection of similar 3D folds between two polypeptides that have been developed. These methods are used to  
20 identify protein structural similarity, also know as similarity of 3D fold. The application of these methods is described collectively as Fold Recognition.

Fold recognition was originally envisioned as a tool to be used for validation of experimentally determined structures or hypothetical models. Prior to 1990, the three-dimensional protein databases contained several 3D structures which were not  
25 recognized as incorrect until a subsequent 3D structure was determined for the same molecule. Several fold recognition methods were developed and automated in the early 1990s including 3D profiles developed by the Eisenberg laboratory (Luthy et al., 1992, Fischer & Eisenberg, 1996) and protein threading as implemented by the Sippl group (Sippl , 1990; Sippl & Weitckus, 1992; Sippl, 1993), Thornton and Jones  
30 (Jones et. al., 1992), Skolnick & Godzik (Godzik, 1992) laboratories.

These fold recognition tools were immediately used for validation of experimentally determined structures in the public Protein Data Bank (PDB,

Bernstein et. al. 1977). It was also recognized that these profile and threading methods could be used to evaluate the quality of molecular models constructed by comparative and homology modeling (Sali et al., Proteins 23:318-26 (1995)). The principle of fold recognition is that a library of known protein structures (derived from the Protein Data Bank, Bernstein et. al. 1977) can be used to find the best 3D template that matches a given query sequence. Fold recognition methods such as protein threading fit a query sequence directly onto the cartesian coordinates of template structures. The plausibility of the molecular model created by the threading of a sequence onto a template three dimensional structure is evaluated by a scoring function constructed from a (statistical) knowledge database of known protein structure. To exploit structural information inherent in the molecular model, functional sites can be evaluated so that the structural models can be used to assign biochemical function. The Rose laboratory ( Xu et al. 1999) made two such successful predictions on proteins from the *Methanococcus* genome that were classified as “hypothetical”. The functional predictions were subsequently confirmed experimentally.

Fold recognition methods have been applied to microbial genome annotation (Fischer & Eisenberg Proc. Natl. Acad. Sci., USA 94:11929,1997; Huynen et. al. 1998; Jones et. al. 1998; Rychlewski et. al. 1998; Rychlewski et. al. 1999; Pawlowski et. al. 1999). For each of the annotation methods cited above, screening of the fold predictions showed that fold predictions could be used to extend functional annotations in a genome. The functional conservation of a gene is based upon conservation of specific residues in active or other functional sites. Results varied depending upon the protein structure library used and the genome annotated. For purposes of genome annotation, novel folds cannot be recognized by fold recognition methods. Current estimates suggest that between 60 and 80 percent of fold domains are known. Experimental methods such as X-ray or NMR spectroscopy can be used to elucidate the three dimensional structures for these novel folds and this information is constantly being added to the public structure database, PDB. Addition of this information will increase the ability of fold recognition methods to detect three dimensional fold similarities.

Correlation of protein sequence with protein structure next brings the “paradox” of structural genomics. To what extent can the function of a protein be

deduced from structure? The correlation of protein structure and function was explored by the Sippl group (Koppensteiner et al. 2000) who showed that proteins that have similar folds usually have similar functions. The study concluded that, even when sequence similarity was low (e.g., less than 30% identity), in 66% of the cases structural similarity corresponded to similarity in function. Published studies on genome annotation that utilize fold recognition technology suggest that, if sequence based methods can annotate 20-40% of a genome with high confidence, an additional 5-20% of the genome can be reliably annotated using fold recognition methods. In summary, it is clear that fold recognition methods can be used to complement the sequenced-based bioinformatics methods (mentioned above) to uncover additional functional relationships for genomic sequences.

### Methodology

The goal of this method was to use the structural information from a family of protein domains to provide enhanced recognition and functional assignment for novel (hypothetical) and incorrectly annotated genome sequences. Developed protocol is similar to the sequence-based profile methods such as PSI-BLAST (Altschul et. al. 1997). It has been shown that the sequence based profile methods can incorporate position specific variation within a given protein family to extend the ability to detect remote sequence and functional conservation. In order to augment mutation information available for specific protein families, a protein threading protocol has been developed that enhances the sequence information by incorporating three dimensional residue interaction preferences for a related family of protein structures. The result provides a structural signature that, when applied across a protein family, allows for the recognition of structural similarity even in the absence of homology.

The protocol is best described as a structural genomics approach to genome mining. Briefly, there are four stages to this protocol for genome mining:

#### ProtocolGenome Sequences

- 1.Template Library Generation
- 2.Genome Wide Fold RecognitionList 1
- 3.Novelty ScanList 1 -> List 2 -> List 3
- 4.*in silico* PredictionList 3 -> List 4

## 1.Template Library Generation

The first stage is the generation of the template library that will be used for fold recognition. This is the most critical stage since protein threading is dependent upon the library of known structures used for template matching. The goal of this stage was to gather the three dimensional structures for a specific protein family or superfamily. Methods which can be used to generate template libraries include but are not limited to:

A) Selection based upon previous structural classification databases available SCOP, CATH, DALI.

B) Selection based upon characterization of single domains that correspond to catalytic or functional domains of interest within the protein family of interest.

C) Selection which incorporates proprietary (in-house) Bristol-Myers Squibb three dimensional structures.

D) Selection of templates from the Protein Data Bank (Bernstein et. al. 1977) using numerous sequence and profile based methods (some of which are available on the PDB web site [www.rcsb.org/pdb](http://www.rcsb.org/pdb)).

E) Selection of templates can also be made using three dimensional structural comparison tools (Zu-Kang & Sippl 1996, Lackner et. al. 2000) to collect and compare related three dimensional structures.

The preferred protocol for stage 1, *Template Library Generation* utilizes a combination of the above mentioned methods which would result in a comprehensive assembly of related protein domains of known structure and related functions.

## 2.Genome Wide Fold Recognition

Stage two consists of threading the entire genome (genome can be defined as a large number of sequences from an organism) onto each structure in the template library. For protein threading, the PROCERYON suite of software (Proceryon Biosciences, New York) was used with the protocol that was automated by programming (perl) scripts. The PROCERYON threading software fits a query sequence directly onto the cartesian coordinates of template structures. The plausibility of the fitted model was evaluated by a scoring function constructed from statistical knowledge based potentials that were derived from proteins of known three-

dimensional structure. The threading scores of interest are the pairwise, surface, combined (pairwise interaction score and the surface score, P/S) energy scores and the sequence identity score (SEQ and ID). The scores for each genome sequence threaded onto each template was stored in a database and an additional score (Threshold Index) calculated. Threshold index is a single score that is a combination of the combined energy score and the sequence (identity score). In order to extract an initial list of genes (List 1) for post threading processing the threshold index score was used in combination with sequence identity and raw sequence length (number of amino acids). A list of genes (List 1) that met the following criteria were selected from this stage two process:

Threading index > 50

Sequence identity > 15%

Sequence length > 100 amino acids and < 700 amino acids

### 3. Novelty Scan

The list of gene sequences (List 1) resulting from the previous stage were processed in stage three using two successive tactics. The first tactic was to “back-thread” the genomic list (List 1) against a template library that contains at least one representative of each protein and functional domain. For this study we used the PDB40 for the template library. The PDB40 is a template library generated from all known protein structures that have less than 40% identity to each other. This tactic was aimed at testing the initial structural assignment used to identify the sequence(s) from the previous two stages. The goal was to determine what structural similarity exists between the query sequence and members of the protein and functional domain templates. The results of the back-threading were compared to the original structural assignment. Sequences were removed from the list (List 1) if more significant structural and functional relationships were found to templates from the protein and functional domain library (PDB40). The sequences remaining on the list (List 2) were passed onto the second tactic.

The second tactic was to perform successive exhaustive BLAST searches on the list (List 2) against a series of sequence databases which may include but are not

limited to the following databases: non-redundant protein sequences, non-redundant nucleotide sequences, ESTs, Incyte Templates (LifeSeq Gold, Incyte Pharmaceuticals, Inc.) and databases of patented genes and protein sequences (GenSeq Database, Derwent, Inc). The result of this stage (Stage 3) was the generation of a list (List 3) of gene sequences for which a putative biological function was determined and the novelty of the sequence assignment assessed by back-threading and exhaustive BLAST searches.

#### 4. *in silico* Prediction

Each sequence in the list (List 3) was analyzed manually for possible functional conservation to members of the protein family template library and related sequences (Stage 1). Preferred procedures for this validation include but are not limited to:

A) Pairwise sequence alignment and conservation of functional site residues.

B) Multiple sequence alignment of members of the protein family template library and/or related protein sequences.

C) Multiple sequence alignments with proteins of known function whose relationships were determined based upon profile search methods such as PSI-BLAST and Hidden Markov Models (eg. Pfam, etc.).

D) Three dimensional homology or comparative modeling where a three dimensional model is used to help validate function.

E) Conservation profiles, sequence or structural motifs used to characterize the functional residues in catalytic, binding, allosteric and other functional sites.

F) Manual adjustment of the sequence alignment where the sequence(s) are aligned “by hand” for comparison to multiple sequence alignments derived from any or all of the steps A-E above.

The results of Stage 4 was a list (List 4) of sequences for which there was structural and or functional characterization based upon the *in silico* protocol.

The partial BMY\_HPP13 (Figure 2) polynucleotide (SEQ ID NO:3) and polypeptide (SEQ ID NO:4) of the present invention was first identified as belonging to the phosphatase family using the above structural threading methods based upon its structural alignment to the human Shp-2 sequence and the human CDC25B sequence (see Figure 3B and 3A; respectively). The partial BMY\_HPP13 sequence was used to

### EXAMPLE 3 – METHOD OF CONVERTING DOUBLE STRANDED cDNA LIBRARIES INTO SINGLE STRAND CIRCULAR FORMS

#### Preparation of culture

LB medium (200 mL+ 400 ul carb) is inoculated with 0.2 to 1 ml of thawed  
5 cDNA library. The culture is incubated, shaking at 250 rpm at 37°C for 45 min. The  
optical density of the culture is measured. The OD600 is preferably between 0.025  
and 0.040. One mL M13K07 helper phage is added to the culture and grown for 2  
hours. At that time, 500 uL Kanamycin (30 mg/mL) is added and incubation  
continued for 15-18 hours.

#### 10 Preparation of cells for precipitation

Cultures are poured into six 50 mL tubes. Cells are centrifuged at 10000 rpm  
in an HB-6 rotor for 15 minutes at 4°C. The supernatant is retrieved and cells  
discarded. The supernatant is filtered through a 0.2 um filter. DNase I (12000 units  
from Gibco) is added and incubated at room temperature for 90 minutes.

#### 15 PEG precipitation of DNA

Fifty mL of ice-cold 40% PEG 8000, 2.5 M NaCl, 10 mM MgSO<sub>4</sub> is added to  
the cell pellets. The solution is mixed and distributed into 6 centrifuge tubes and  
covered with parafilm. The tubes are incubated on wet ice for 1 hour (or at 4°C  
overnight).

20 Phage are pelleted at 10000 rpm in an HB-6 rotor for 20 minutes at 4°C. The  
supernatant is discarded and the sides of the tubes wiped dry. The pellets are  
resuspended in 1mL TE, pH 8.

The resuspended pellets are placed in a 14 mL Sarstedt tube (6 mL total). SDS  
is added to 0.1% (60 uL of stock 10% SDS). Proteinase K (60 uL of 20 mg/mL) is  
25 then added and incubated at 42C for 1 hour.

DNA is extracted with phenol/chloroform by first adding 1 mL of 5M NaCl  
followed by an equal volume of phenol/chloroform (6 mL). The mixture is vortexed  
and centrifuged at 5K in an HB-6 rotor for 5 minutes at 4°C. The aqueous (top) phase  
is transferred to a new Sarstedt tube. Extractions are repeated until no interface is  
30 visible.

The DNA is precipitated in ethanol by adding 2 volumes of 100% ethanol and  
precipitating overnight at -20°C. The DNA is centrifuged at 10000 rpm in HB-6 rotor

BLAST against the human genome database. This resulted in the identification of human BAC AC06831 as the portion of the genome harboring this gene. The GENEWISEDB algorithm was then applied to the BAC AC06831 sequence to elucidate the exon/intron structural of the BMY\_HPP13 gene (see Figure 4).

5        Appropriate primers were designed based upon the genomic structure of the BMY\_HPP13 gene and the full-length clone was isolated as described herein. The full-length sequence of the BMY\_HPP13 polynucleotide is provided in Figures 1A-B (SEQ ID NO:1)

## 10        EXAMPLE 2 – METHOD OF CONSTRUCTING A SIZE FRACTIONATED              BRAIN AND TESTIS cDNA LIBRARY

Poly A<sup>+</sup> RNA from Clontech is treated with DNase I to remove genomic DNA contamination. The RNA is converted into double stranded cDNA using the SuperScript<sup>TM</sup> Plasmid System for cDNA Synthesis and Plasmid Cloning (Life  
15   Technologies). The cDNA is size fractionated on a TransGenomics HPLC size exclusion column (TosoHass) with dimensions of 7.8mm x 30cm and a particle size of 10µm. Tris buffered saline is used as the mobile phase, and the column is run at a flow rate of 0.5 ml/min. The system is calibrated using a 1 kb ladder to determine which fractions are to be pooled to obtain the largest cDNA library. Generally,  
20   fractions that eluted in the range of 12 to 15 minutes are used. The cDNA is precipitated, concentrated and then ligated into the Sal I / Not I sites in pSPORT. Following electroporation of the cDNA into DH12S, DNA from the resulting colonies is prepared and subjected to Sal I/Not I restriction enzyme digestion. Generally, the average insert size of libraries made by this procedure is greater than 3.5 Kb and the  
25   overall complexity of the library is greater than 10<sup>7</sup> independent clones. The library is amplified in semi-solid agar for 2 days at 30 C. An aliquot (200 microliters) of the amplified library is inoculated into a 200 ml culture for single-stranded DNA isolation by super-infection with a fi helper phage. The single-stranded circular DNA is concentrated by ethanol precipitation, resuspended at a concentration of one  
30   microgram per microliter and used for the cDNA capture experiments.

for 20 minutes at 4°C. The ethanol is discarded and the pellets resuspended in 700 uL 70% ethanol. The resuspended pellets are centrifuged at 14000 rpm for 10 minutes at 4°C. The ethanol is discarded and the pellets dried by vacuum.

Oligosaccharides are then removed by resuspending the pellet in 50 uL TE, pH 8. The solutions are frozen on dry ice for 10 minutes and centrifuged at 14000rpm for 15 minutes at 4°C. The supernatant is transferred to a new tube and the volume recorded.

The concentration of DNA is determined by measuring absorbance at 260/280. DNA is diluted 1:100 in a quartz cuvette (3 uL DNA + 297 uL TE). The following equation is used to calculate DNA concentration:

$$(32 \text{ ug/mL} \cdot \text{OD})(\text{mL}/100 \text{ uL})(100)(\text{OD}_{260}) = \text{DNA concentration}$$

The preferred purity ratio is 1.7 - 2.0.

The DNA is diluted to 1 ug/uL with TB, pH 8 and stored at 4°C.

To test the quality of single-stranded DNA (ssDNA) the following reaction mixtures are prepared:

1. DNA mix per reaction

- a. 1 uL of 5 ng/uL ssDNA (1:200 dilution of VI.D.2 above)
- b. 11 uL dH<sub>2</sub>O
- c. 1.5 uL 10 uM T7 SPORT primer (fresh dilution of stock)
- d. 1.5 uL 10X Precision-Taq buffer

2. Repair mix per reaction

- a. 4 uL 5 mM dNTPs (1.25 mM each)
- b. 1.5 uL 10X Precision-Taq buffer
- c. 9.25 uL dH<sub>2</sub>O
- d. 0.25 uL Precision-Taq polymerase
- e. Preheat cocktail at 70°C until middle of thermal cycle

The DNA mixes are aliquoted into PCR tubes and thermal cycle carried out as follows:

- 1. 95°C, 20 sec
- 2. 59°C, 1 min; add 15 uL repair mix
- 3. 73°C, 23 min

Ethanol precipitation of the ssDNA is performed by adding 15 ug glycogen, 16 uL 7.5 M NH<sub>4</sub>OAc, 125 uL 100% ethanol. The sample is centrifuged at 14000 rpm for 30 minutes at 4°C and the pellet washed with 125 uL 70% ethanol. The ethanol is discarded and pellet dried by vacuum. The pellet is resuspended in 10 uL TB, pH 8.

- 5           The DNA is electroporated into DH10B or DH12S cells. A DNA mixture consisting of:

1.2 uL repaired library (=1.0 x 10<sup>-3</sup> ug)

2.1 uL 1 ng/uL unrepaired library (= 1.0 x 10<sup>-3</sup> ug)

3.1 uL 0.01 ug/uL pUC19 positive control DNA (= 1 x 10<sup>-5</sup> ug)

- 10       is aliquoted to Eppendorf tubes. Cells are thawed on ice-water. Forty uL of cells are added to each DNA aliquot by pipetting into a chilled cuvette placed between metal plates. Electroporation is carried out at 1.8 kV. Immediately following electroporation, 1 mL SOC (SOB + glucose + Mg<sup>++</sup>) media is added to the cuvette, then transferred to a 15 mL tube. Cells are allowed to recover for 1 hr at 37°C with
- 15       shaking (225 rpm). Cells are then plated according to the following dilution scheme:

#### A. Dilutions of Culture

1. Serial dilutions of culture in 1:10 increments (20 uL into 180 uL LB broth)

2. Repaired dilutions

a. 1:100

20       b. 1:1K

c. 1:10K

3. Unrepaired dilutions

a. 1:10

b. 1:100

25       4. Positive control dilutions

a. 1:10

b. 1:100

100 uL of each dilution is plated on small LB+carb plates and incubated at 37°C overnight. Colonies are counted to calculate titer as follows:

- 30       1. use smallest countable dilution
2. (# of colonies)(dilution factor)(200 uL/100 uL)(1000 uL/20 uL) = CFUs
3. CFUs / ug DNA used = CFU/ug

% Background = (unrepaired CFU/ug / repaired CFU/ug) x 100%

#### EXAMPLE 4 – METHOD OF CLONING THE NOVEL HUMAN BMY\_HPP13 POLYPEPTIDE OF THE PRESENT INVENTION

5           One microliter of anti-sense biotinylated oligos (or sense oligos when annealing to single stranded DNA from pSPORT2 vector), containing one hundred and fifty nanograms of 1 to 50 different 80mer oligo probes, is added to six microliters (six micrograms) of a mixture of up to 15 single-stranded covalently closed circular cDNA libraries and seven microliters of 100% formamide in a 0.5 ml  
10   PCR tube. The sequence of the 80mer oligo used is as follows:

5'-TAAAGACACAGATGTTTCAGTGGATCTGGGTCTCGACTGGGCCCTAATT  
TCTCATACCCACTCCCCTTAGCCTCTTTTGCC-3' (SEQ ID NO:28).

15           The mixture is heated in a thermal cycler to 95°C for 2 min. Fourteen microliters of 2X hybridization buffer (50% formamide, 1.5 M NaCl, 0.04 M NaPO<sub>4</sub>, pH 7.2, 5 min EDTA, 0.2% SDS) is added to the heated probe/cDNA library mixture and incubated at 42°C for 26 hours. Hybrids between the biotinylated oligo and the circular cDNA are isolated by diluting the hybridization mixture to 220 microliters  
20   solution containing 1 M NaCl, 10 mM Tris-HCl pH 7.5, 1mM EDTA, pH 8.0 and adding 125 microliters of streptavidin magnetic beads. This solution is incubated at 42°C for 60 min, and mixed every 5 min to re-suspend the beads. The beads are separated from the solution with a magnet and washed three times in 200 microliters of 0.1 X SSPE, 0.1% SDS at 45°C.

25           The single stranded cDNA is released from the biotinylated oligo/streptavidin magnetic bead complex by adding 50 microliters of 0.1 N NaOH and incubating at room temperature for 10 min. Six microliters of 3 M sodium acetate is added along with 15 micrograms of glycogen and the solution ethanol precipitated with 120 microliters of 100% ethanol. The precipitated DNA is resuspended in 12 microliters  
30   of TB (10 mM TrisHCl, pH 8.0), 1mM EDTA, pH 8.0). The single-stranded cDNA is converted into double-stranded DNA in a thermal cycler by mixing 5 microliters of the captured DNA with 1.5 microliters of 10 micromolar standard SP6 primer for

libraries in pSPORT 1 and 2 and 17 primer for libraries in pCMVSPORT and 1.5 microliters of 10 X PCR buffer.

Sequences of primers used to repair single-stranded circular DNA isolated from the primary selection are as follows:

5

T7Sport5'-TAATACGACTCACTATAGGG-3' (SEQ ID NO:58)

SP6Sport5'- ATTTAGGTGACACTATAG -3' (SEQ ID NO:59)

10 The mixture is heated to 95°C for 20 seconds and the temperature gradually brought down to 59°C. Fifteen microliters of a repair mix, that was preheated to 70°C is added to the DNA (repair mix contains 4 microliters of 5 mM dNTPs (1.25 mM each), 1.5 microliters of 10X PCR buffer, 9.25 microliters of water, and 0.25 microliters of Taq polymerase). The solution incubation temperature is raised back to 73°C and incubated for 23 min. The repaired DNA is ethanol precipitated and  
15 resuspended in 10 microliters of TB. Electroporation is carried out using two microliters DNA per 40 microliters of E. coli DH12S cells. Three hundred and thirty three microliters are plated onto one 150-mm plate of LB agar plus 100 micrograms/milliliter of ampicillin. After overnight incubation at 37°C, the colonies from all plates are harvested by scraping into 10 ml of LB medium +  
20 50 micrograms/milliliter of ampicillin and 2 ml of sterile glycerol.

The second round of selection is initiated by making single-stranded circular DNA from the primary selected library using the method listed above. The purified single-stranded circular DNA is then assayed with gene-specific primers for each of the targeted sequences using standard PCR conditions.

25 The sequences of the Gene-Specific-Primer ("GSP") pairs used to identify the various targeted cDNAs in the primary selected single stranded cDNA libraries are as follows:

Left Primer 1:TCCCAATATGAGATGCCTGA (SEQ ID NO:31)

30 Right Primer 1:AGCTGACTGGTTCTTGGCTT (SEQ ID NO:32)

The secondary hybridization is carried out using only those 80mer biotinylated probes whose targeted sequences were positive with the GSPs. The resulting single-stranded circular DNA is converted to double strands using the antisense oligo for each target sequence as the repair primer (the sense primer is used for material  
 5 captured from pSPORT2 libraries. The resulting double stranded DNA is electroporated into DH10B and the resulting colonies inoculated into 96 deep well blocks. Following overnight growth, DNA is prepared and sequentially screened for each of the targeted sequences using the GSPs. The DNA is also cut with Sal I and Not I and the inserts sized by agarose gel electrophoresis.

10 Those cDNA clones that were positive by PCR had the inserts sized and two clones were chosen for DNA sequencing for each gene. All of the clones had identical

#### EXAMPLE 5 – METHOD OF ASSESSING THE EXPRESSION PROFILE OF THE NOVEL BMY\_HPP13 POLYPEPTIDES OF THE PRESENT INVENTION USING 15 EXPANDED mRNA TISSUE AND CELL SOURCES

Total RNA from tissues was isolated using the TriZol protocol (Invitrogen) and quantified by determining its absorbance at 260nm. An assessment of the 18s and 28s ribosomal RNA bands was made by denaturing gel electrophoresis to determine RNA integrity.

20 The specific sequence to be measured was aligned with related genes found in GenBank to identify regions of significant sequence divergence to maximize primer and probe specificity. Gene-specific primers and probes were designed using the ABI primer express software to amplify small amplicons (150 base pairs or less) to maximize the likelihood that the primers function at 100% efficiency. All  
 25 primer/probe sequences were searched against Public Genbank databases to ensure target specificity. Primers and probes were obtained from ABI.

For BMY\_HPP13, the primer probe sequences were as follows:

Forward Primer 5'-TCAAGGGTGGGAAGCAATACCA -3' (SEQ ID NO:13)

30 Reverse Primer 5'-CTTCGCTGGCAGGAGGAA -3' (SEQ ID NO:14)

TaqMan Probe 5' -CCGGAACCAGCCAAATGCTCTCTG -3' (SEQ ID NO:15)

### DNA contamination

To access the level of contaminating genomic DNA in the RNA, the RNA was divided into 2 aliquots and one half was treated with Rnase-free Dnase (Invitrogen). Samples from both the Dnase-treated and non-treated were then subjected to reverse transcription reactions with (RT+) and without (RT-) the presence of reverse transcriptase. TaqMan assays were carried out with gene-specific primers (see above) and the contribution of genomic DNA to the signal detected was evaluated by comparing the threshold cycles obtained with the RT+/RT- non-Dnase treated RNA to that on the RT+/RT- Dnase treated RNA. The amount of signal contributed by genomic DNA in the Dnased RT- RNA must be less than 10% of that obtained with Dnased RT+ RNA. If not the RNA was not used in actual experiments.

### Reverse Transcription reaction and Sequence Detection

100ng of Dnase-treated total RNA was annealed to 2.5  $\mu$ M of the respective gene-specific reverse primer in the presence of 5.5 mM Magnesium Chloride by heating the sample to 72°C for 2 min and then cooling to 55° C for 30 min. 1.25 U/ $\mu$ l of MuLv reverse transcriptase and 500 $\mu$ M of each dNTP was added to the reaction and the tube was incubated at 37° C for 30 min. The sample was then heated to 90°C for 5 min to denature enzyme.

Quantitative sequence detection was carried out on an ABI PRISM 7700 by adding to the reverse transcribed reaction 2.5 $\mu$ M forward and reverse primers, 500 $\mu$ M of each dNTP, buffer and 5U AmpliTaq Gold™. The PCR reaction was then held at 94°C for 12 min, followed by 40 cycles of 94° C for 15 sec and 60° C for 30 sec.

### Data handling

The threshold cycle (Ct) of the lowest expressing tissue (the highest Ct value) was used as the baseline of expression and all other tissues were expressed as the relative abundance to that tissue by calculating the difference in Ct value between the baseline and the other tissues and using it as the exponent in  $2^{(\Delta Ct)}$

The expression profile of the BMY\_HPP13 polypeptide is provided in Figure 5 and described elsewhere herein.

EXAMPLE 6 – METHOD OF ASSAYING THE PHOSPHATASE ACTIVITY OF  
THE BMY\_HPP POLYPEPTIDES OF THE PRESENT INVENTION

The Phosphatase Activity of the BMY\_HPP Polypeptides of the present invention may be assessed through the application of various biochemical assays  
5 known in the art and described herein.

Hydrolysis of para-nitrophenyl phosphate

The phosphatase activity for BMY\_HPP proteins may be measured by assaying the ability of the proteins to hydrolyze para-nitrophenyl phosphate, a compound known to be a substrate for phosphatases, as described in Krejsa, C. et al.,  
10 J. Biol. Chem. Vol. 272, p.11541-11549, 1997 (which is hereby incorporated in its entirety herein). The proteins are incubated in 3 mg/ml para-nitrophenyl phosphate in a solution containing 60 mM MES, pH 6.0, 5% glycerol, 5 mM dithiothreitol, and 0.1% Triton X-100 for 15 min, or such other time as may be desired. The pH of the reaction may be varied to provide an optimal pH for each individual BMY\_HPP  
15 protein by those with ordinary skill in the art of enzyme assays. The phosphatase reaction is stopped by the addition of 3 N NaOH to give a final NaOH concentration of 0.7 M. The product of the reaction is measured by reading the absorbance of the solution at 405 nm.

Two dimensional gel electrophoresis

20 The BMY\_HPP polynucleotides of the present invention may be subcloned into appropriate vectors for expression in host cells. Representative vectors are known in the art and described herein. 2-D gel electrophoresis (IEF followed by SDS-PAGE) will be used to assay BMY\_HPP-dependent dephosphorylation of host cell proteins. These proteins can be recovered from the gel and identified by mass spectrometric or  
25 other protein sequencing techniques known in the art.

Briefly, Methods for 2-dimensional gel analysis and labeling cells with proteins are well known in the art. Cells would be labeled with  $^{32}\text{P}$  orthophosphate, cellular proteins would be resolved on 2D gels and their positions determined by autoradiography. Proteins of interest would be identified by excising the spots and  
30 analyzing their sequence by mass spectroscopy. The following paper and the references therein describe the methods of labeling cells, analyzing the proteins on 2D gels and mass spec identification: Gerner, C. et al., J. Biol. Chem., Vol. 275, p.39018-

39026, 2000. Substrates affected by the phosphatase would be identified by comparing wild type cells to cells where expression of the phosphatase is inhibited by deletion, anti-sense, or other means. Proteins whose phosphorylation increased would be either direct substrates or indirectly regulated by the phosphatase. Conversely, in  
5 cells where the active phosphatase was overexpressed, proteins whose phosphorylation decreased would either be direct substrates or indirectly regulated by the phosphatase.

EXAMPLE 7 – METHOD OF IDENTIFYING THE SUBSTRATE OF THE  
10 BMY\_HPP PHOSPHATASE POLYPEPTIDES OF THE PRESENT INVENTION  
Substrate identification

Protein substrates for BMY\_HPP polypeptides of the present invention may be identified by recovery of proteins dephosphorylated in the 2-D gel electrophoresis assay described above. Phosphopeptide substrates may also be identified as proteins  
15 whose phosphorylation increases when the activity or expression of a BMY\_HPP protein is decreased (for example, by an antibody, antisense or double-stranded inhibitory RNA or by a small molecule inhibitor of BMY\_HPP activity). In either case, mass spectrometry can be used to identify the recovered proteins.

Phosphopeptide substrates for BMY-HPP polypeptides may also be identified  
20 by incubation of a phosphopeptide library with a catalytically inactive version of the protein, recovery of the complex, and peptide sequencing by standard methods such as Edman degradation or mass spectrometry.

Phosphopeptide substrates can also be identified by expressing a substrate trapping mutant phosphatase (one that is catalytically inactive due to active site  
25 mutation) and isolating the proteins that bind preferentially to the substrate trapping phosphatase relative to the wild type phosphatase.

EXAMPLE 8 – METHOD OF ASSESSING THE PHYSIOLOGICAL FUNCTION  
OF THE HUMAN PHOSPHATASE POLYPEPTIDE AT THE CELLULAR LEVEL

30 The physiological function of the human phosphatase polypeptide may be assessed by expressing the sequences encoding human phosphatase at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a

mammalian expression vector containing a strong promoter that drives high levels of cDNA expression (examples are provided elsewhere herein). Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10, ug of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2ug of an additional plasmid containing sequences encoding a marker protein are cotransfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M. G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of human phosphatase polypeptides on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding human phosphatase and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding

human phosphatase polypeptides and other genes of interest can be analyzed by northern analysis or microarray techniques.

EXAMPLE 9 – METHOD OF SCREENING FOR COMPOUNDS THAT  
INTERACT WITH THE HUMAN PHOSPHATASE POLYPEPTIDE

The following assays are designed to identify compounds that bind to the human phosphatase polypeptide, bind to other cellular proteins that interact with the human phosphatase polypeptide, and to compounds that interfere with the interaction of the human phosphatase polypeptide with other cellular proteins.

Such compounds can include, but are not limited to, other cellular proteins. Specifically, such compounds can include, but are not limited to, peptides, such as, for example, soluble peptides, including, but not limited to Ig-tailed fusion peptides, comprising extracellular portions of human phosphatase polypeptide transmembrane receptors, and members of random peptide libraries (see, e.g., Lam, K. S. et al., 1991, Nature 354:82-84; Houghton, R. et al., 1991, Nature 354:84-86), made of D-and/or L-configuration amino acids, phosphopeptides (including, but not limited to, members of random or partially degenerate phosphopeptide libraries; see, e.g., Songyang, Z., et al., 1993, Cell 72:767-778), antibodies (including, but not limited to, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and FAb, F(ab').sub.2 and FAb expression library fragments, and epitope-binding fragments thereof), and small organic or inorganic molecules.

Compounds identified via assays such as those described herein can be useful, for example, in elaborating the biological function of the human phosphatase polypeptide, and for ameliorating symptoms of tumor progression, for example. In instances, for example, whereby a tumor progression state or disorder results from a lower overall level of human phosphatase expression, human phosphatase polypeptide, and/or human phosphatase polypeptide activity in a cell involved in the tumor progression state or disorder, compounds that interact with the human phosphatase polypeptide can include ones which accentuate or amplify the activity of the bound human phosphatase polypeptide. Such compounds would bring about an effective increase in the level of human phosphatase polypeptide activity, thus ameliorating symptoms of the tumor progression disorder or state. In instances

whereby mutations within the human phosphatase polypeptide cause aberrant human phosphatase polypeptides to be made which have a deleterious effect that leads to tumor progression, compounds that bind human phosphatase polypeptide can be identified that inhibit the activity of the bound human phosphatase polypeptide.

- 5 Assays for testing the effectiveness of such compounds are known in the art and discussed, elsewhere herein.

#### EXAMPLE 10 – METHOD OF SCREENING, IN VITRO, COMPOUNDS THAT BIND TO THE HUMAN PHOSPHATASE POLYPEPTIDE

- 10 In vitro systems can be designed to identify compounds capable of binding the human phosphatase polypeptide of the invention. Compounds identified can be useful, for example, in modulating the activity of wild type and/or mutant human phosphatase polypeptide, preferably mutant human phosphatase polypeptide, can be useful in elaborating the biological function of the human phosphatase polypeptide,  
15 can be utilized in screens for identifying compounds that disrupt normal human phosphatase polypeptide interactions, or can in themselves disrupt such interactions.

- The principle of the assays used to identify compounds that bind to the human phosphatase polypeptide involves preparing a reaction mixture of the human phosphatase polypeptide and the test compound under conditions and for a time  
20 sufficient to allow the two components to interact and bind, thus forming a complex which can be removed and/or detected in the reaction mixture. These assays can be conducted in a variety of ways. For example, one method to conduct such an assay would involve anchoring human phosphatase polypeptide or the test substance onto a solid phase and detecting human phosphatase polypeptide /test compound complexes  
25 anchored on the solid phase at the end of the reaction. In one embodiment of such a method, the human phosphatase polypeptide can be anchored onto a solid surface, and the test compound, which is not anchored, can be labeled, either directly or indirectly.

- In practice, microtitre plates can conveniently be utilized as the solid phase. The anchored component can be immobilized by non-covalent or covalent  
30 attachments. Non-covalent attachment can be accomplished by simply coating the solid surface with a solution of the protein and drying. Alternatively, an immobilized antibody, preferably a monoclonal antibody, specific for the protein to be

immobilized can be used to anchor the protein to the solid surface. The surfaces can be prepared in advance and stored.

In order to conduct the assay, the nonimmobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (e.g., by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the previously immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously nonimmobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the immobilized component (the antibody, in turn, can be directly labeled or indirectly labeled with a labeled anti-Ig antibody).

Alternatively, a reaction can be conducted in a liquid phase, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for human phosphatase polypeptide or the test compound to anchor any complexes formed in solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

20

#### EXAMPLE 11 – METHOD OF IDENTIFYING COMPOUNDS THAT INTERFERE WITH HUMAN PHOSPHATASE POLYPEPTIDE/CELLULAR PRODUCT INTERACTION

The human phosphatase polypeptide of the invention can, in vivo, interact with one or more cellular or extracellular macromolecules, such as proteins. Such macromolecules include, but are not limited to, polypeptides, particularly ligands, and those products identified via screening methods described, elsewhere herein. For the purposes of this discussion, such cellular and extracellular macromolecules are referred to herein as "binding partner(s)". For the purpose of the present invention, "binding partner" may also encompass polypeptides, small molecule compounds, polysaccharides, lipids, and any other molecule or molecule type referenced herein. Compounds that disrupt such interactions can be useful in regulating the activity of

25

30

the human phosphatase polypeptide, especially mutant human phosphatase polypeptide. Such compounds can include, but are not limited to molecules such as antibodies, peptides, and the like described in elsewhere herein.

5       The basic principle of the assay systems used to identify compounds that interfere with the interaction between the human phosphatase polypeptide and its cellular or extracellular binding partner or partners involves preparing a reaction mixture containing the human phosphatase polypeptide, and the binding partner under conditions and for a time sufficient to allow the two products to interact and bind, thus forming a complex. In order to test a compound for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound can be initially included in the reaction mixture, or can be added at a time subsequent to the addition of human phosphatase polypeptide and its cellular or extracellular binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the human phosphatase polypeptide and the cellular or extracellular binding partner is then detected. The formation of a complex in the control reaction, but not in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the human phosphatase polypeptide and the interactive binding partner. Additionally, complex formation within reaction mixtures containing the test compound and normal human phosphatase polypeptide can also be compared to complex formation within reaction mixtures containing the test compound and mutant human phosphatase polypeptide. This comparison can be important in those cases wherein it is desirable to identify compounds that disrupt interactions of mutant but not normal human phosphatase polypeptide.

25       The assay for compounds that interfere with the interaction of the human phosphatase polypeptide and binding partners can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the human phosphatase polypeptide or the binding partner onto a solid phase and detecting complexes anchored on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the

interaction between the human phosphatase polypeptide and the binding partners, e.g., by competition, can be identified by conducting the reaction in the presence of the test substance; i.e., by adding the test substance to the reaction mixture prior to or simultaneously with the human phosphatase polypeptide and interactive cellular or extracellular binding partner. Alternatively, test compounds that disrupt preformed complexes, e.g. compounds with higher binding constants that displace one of the components from the complex, can be tested by adding the test compound to the reaction mixture after complexes have been formed. The various formats are described briefly below.

10 In a heterogeneous assay system, either the human phosphatase polypeptide or the interactive cellular or extracellular binding partner, is anchored onto a solid surface, while the non-anchored species is labeled, either directly or indirectly. In practice, microtitre plates are conveniently utilized. The anchored species can be immobilized by non-covalent or covalent attachments. Non-covalent attachment can be accomplished simply by coating the solid surface with a solution of the human phosphatase polypeptide or binding partner and drying. Alternatively, an immobilized antibody specific for the species to be anchored can be used to anchor the species to the solid surface. The surfaces can be prepared in advance and stored.

20 In order to conduct the assay, the partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted components are removed (e.g., by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, can be directly labeled or indirectly labeled with a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds which inhibit complex formation or which disrupt preformed complexes can be detected.

Alternatively, the reaction can be conducted in a liquid phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a  
5 labeled antibody specific for the other partner to detect anchored complexes. Again, depending upon the order of addition of reactants to the liquid phase, test compounds which inhibit complex or which disrupt preformed complexes can be identified.

In an alternate embodiment of the invention, a homogeneous assay can be used. In this approach, a preformed complex of the human phosphatase polypeptide  
10 and the interactive cellular or extracellular binding partner product is prepared in which either the human phosphatase polypeptide or their binding partners are labeled, but the signal generated by the label is quenched due to complex formation (see, e.g., U.S. Pat. No. 4,109,496 by Rubenstein which utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one  
15 of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances which disrupt human phosphatase polypeptide -cellular or extracellular binding partner interaction can be identified.

In a particular embodiment, the human phosphatase polypeptide can be prepared for immobilization using recombinant DNA techniques known in the art. For  
20 example, the human phosphatase polypeptide coding region can be fused to a glutathione-S-transferase (GST) gene using a fusion vector such as pGEX-5X-1, in such a manner that its binding activity is maintained in the resulting fusion product. The interactive cellular or extracellular product can be purified and used to raise a monoclonal antibody, using methods routinely practiced in the art and described  
25 above. This antibody can be labeled with the radioactive isotope <sup>125</sup>I, for example, by methods routinely practiced in the art. In a heterogeneous assay, e.g., the GST- human phosphatase polypeptide fusion product can be anchored to glutathione-agarose beads. The interactive cellular or extracellular binding partner product can then be added in the presence or absence of the test compound in a manner that allows  
30 interaction and binding to occur. At the end of the reaction period, unbound material can be washed away, and the labeled monoclonal antibody can be added to the system and allowed to bind to the complexed components. The interaction between the

human phosphatase polypeptide and the interactive cellular or extracellular binding partner can be detected by measuring the amount of radioactivity that remains associated with the glutathione-agarose beads. A successful inhibition of the interaction by the test compound will result in a decrease in measured radioactivity.

5           Alternatively, the GST- human phosphatase polypeptide fusion product and the interactive cellular or extracellular binding partner product can be mixed together in liquid in the absence of the solid glutathione-agarose beads. The test compound can be added either during or after the binding partners are allowed to interact. This mixture can then be added to the glutathione-agarose beads and unbound material is  
10           washed away. Again the extent of inhibition of the binding partner interaction can be detected by adding the labeled antibody and measuring the radioactivity associated with the beads.

          In another embodiment of the invention, these same techniques can be employed using peptide fragments that correspond to the binding domains of the  
15           human phosphatase polypeptide product and the interactive cellular or extracellular binding partner (in case where the binding partner is a product), in place of one or both of the full length products.

          Any number of methods routinely practiced in the art can be used to identify and isolate the protein's binding site. These methods include, but are not limited to,  
20           mutagenesis of one of the genes encoding one of the products and screening for disruption of binding in a co-immunoprecipitation assay. Compensating mutations in the gene encoding the second species in the complex can be selected. Sequence analysis of the genes encoding the respective products will reveal the mutations that correspond to the region of the product involved in interactive binding. Alternatively,  
25           one product can be anchored to a solid surface using methods described in this Section above, and allowed to interact with and bind to its labeled binding partner, which has been treated with a proteolytic enzyme, such as trypsin. After washing, a short, labeled peptide comprising the binding domain can remain associated with the solid material, which can be isolated and identified by amino acid sequencing. Also,  
30           once the gene coding for the cellular or extracellular binding partner product is obtained, short gene segments can be engineered to express peptide fragments of the product, which can then be tested for binding activity and purified or synthesized.

EXAMPLE 12 – ISOLATION OF A SPECIFIC CLONE  
FROM THE DEPOSITED SAMPLE

The deposited material in the sample assigned the ATCC Deposit Number  
5 cited in Table I for any given cDNA clone also may contain one or more additional  
plasmids, each comprising a cDNA clone different from that given clone. Thus,  
deposits sharing the same ATCC Deposit Number contain at least a plasmid for each  
cDNA clone identified in Table I. Typically, each ATCC deposit sample cited in  
Table I comprises a mixture of approximately equal amounts (by weight) of about 1-  
10 10 plasmid DNAs, each containing a different cDNA clone and/or partial cDNA  
clone; but such a deposit sample may include plasmids for more or less than 2 cDNA  
clones.

Two approaches can be used to isolate a particular clone from the deposited  
sample of plasmid DNA(s) cited for that clone in Table I. First, a plasmid is directly  
15 isolated by screening the clones using a polynucleotide probe corresponding to SEQ  
ID NO:1.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized  
using an Applied Biosystems DNA synthesizer according to the sequence reported.  
The oligonucleotide is labeled, for instance, with  $^{32}\text{P}$ -(-ATP using T4 polynucleotide  
20 kinase and purified according to routine methods. (E.g., Maniatis et al., *Molecular  
Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring, NY (1982).)  
The plasmid mixture is transformed into a suitable host, as indicated above (such as  
XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as  
those provided by the vector supplier or in related publications or patents cited above.  
25 The transformants are plated on 1.5% agar plates (containing the appropriate selection  
agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate.  
These plates are screened using Nylon membranes according to routine methods for  
bacterial colony screening (e.g., Sambrook et al., *Molecular Cloning: A Laboratory  
Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to  
30 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the  
SEQ ID NO:1 (i.e., within the region of SEQ ID NO:1 bounded by the 5' NT and the

3' NT of the clone defined in Table I) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 ul of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is

5 1.5-5 mM MgCl<sub>2</sub>, 0.01% (w/v) gelatin, 20 uM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94 degree C for 1 min; annealing at 55 degree C for 1 min; elongation at 72 degree C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal

10 band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

#### EXAMPLE 13 – BACTERIAL EXPRESSION OF A POLYPEPTIDE

A polynucleotide encoding a polypeptide of the present invention is amplified

15 using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 12, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the

20 restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Ampr), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

25 The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, that expresses the lacI repressor and also confers kanamycin resistance (Kanr).

30 Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.600) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4 degree C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., supra). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., supra).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4 degree C or frozen at -80 degree C.

#### EXAMPLE 14 – PURIFICATION OF A POLYPEPTIDE FROM AN INCLUSION BODY

The following alternative method can be used to purify a polypeptide expressed in E coli when it is present in the form of inclusion bodies. Unless  
5 otherwise specified, all of the following steps are conducted at 4-10 degree C.

Upon completion of the production phase of the E. coli fermentation, the cell culture is cooled to 4-10 degree C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an  
10 appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate  
15 is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the  
20 pellet is discarded and the polypeptide containing supernatant is incubated at 4 degree C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA  
25 by vigorous stirring. The refolded diluted protein solution is kept at 4 degree C without mixing for 12 hours prior to further purification steps.

To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 um membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The  
30 filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perceptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a

stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of  
5 tandem columns of strong anion (Poros HQ-50, Perceptive Biosystems) and weak anion (Poros CM-20, Perceptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a  
10 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A280 monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be  
15 observed from Coomassie blue stained 16% SDS-PAGE gel when 5 ug of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

## 20           EXAMPLE 15 – CLONING AND EXPRESSION OF A POLYPEPTIDE                   IN A BACULOVIRUS EXPRESSION SYSTEM

In this example, the plasmid shuttle vector pAc373 is used to insert a polynucleotide into a baculovirus to express a polypeptide. A typical baculovirus expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction  
25 sites, which may include, for example BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is often used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak *Drosophila* promoter in  
30 the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated

homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors can be used in place of the vector above, such as pVL941 and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., *Virology* 170:31-39 (1989).

A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 12, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites at the 5' end of the primers in order to clone the amplified product into the expression vector. Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified elsewhere herein (if applicable), is amplified using the PCR protocol described in Example 12. If the naturally occurring signal sequence is used to produce the protein, the vector used does not need a second signal peptide. Alternatively, the vector can be modified to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation

mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

5           Five ug of a plasmid containing the polynucleotide is co-transformed with 1.0 ug of a commercially available linearized baculovirus DNA ("BaculoGoldtm baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One ug of BaculoGoldtm virus DNA and 5ug of the plasmid are mixed in a sterile well of a  
10   microtiter plate containing 50ul of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 ul Lipofectin plus 90 ul Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is  
15   then incubated for 5 hours at 27 degrees C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27 degrees C for four days.

          After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, supra. An agarose gel with "Blue Gal" (Life  
20   Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a  
25   micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 ul of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4 degree C.

30           To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection

("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 uCi of 35S-methionine and 5 uCi 35S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

#### EXAMPLE 16 – EXPRESSION OF A POLYPEPTIDE IN MAMMALIAN CELLS

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used include, human HeLa, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transformation with a

selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transformed cells.

The transformed gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

A polynucleotide of the present invention is amplified according to the protocol outlined in herein. If the naturally occurring signal sequence is used to produce the protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.) The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("GeneClean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. E. coli HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for transformation. Five  $\mu$ g of an expression plasmid is cotransformed with 0.5  $\mu$ g of the plasmid pSVneo using lipofectin (Felgner et al., supra). The plasmid pSV2-neo contains a dominant selectable marker, the neo gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded

in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 uM, 2 uM, 5 uM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 uM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

#### EXAMPLE 17 – PROTEIN FUSIONS

The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, and maltose binding protein facilitates purification. (See Example described herein; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the half-life time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule.

Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

The naturally occurring signal sequence may be used to produce the protein (if applicable). Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891 and/or US Patent No. 6,066,781, *supra*.)

5 Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAAACCTCACACATGCCCACCGTGC  
 CCAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCAAAA  
 CCAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTACATGCGTGGT  
 GGTGGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGG  
 10 ACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTA  
 CAACAGCACGTACCGTGTGGTCAGCGTCCTACCGTCCTGCACCAGGACT  
 GGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCA  
 ACCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC  
 CACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAG  
 15 GTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGT  
 GGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCT  
 CCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTG  
 GACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCA  
 TGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG  
 20 GTAAATGAGTGCGACGGCCGCGACTCTAGAGGAT (SEQ ID NO:34)

EXAMPLE 19 – REGULATION OF PROTEIN EXPRESSION VIA  
 CONTROLLED AGGREGATION IN THE ENDOPLASMIC RETICULUM

As described more particularly herein, proteins regulate diverse cellular  
 25 processes in higher organisms, ranging from rapid metabolic changes to growth and  
 differentiation. Increased production of specific proteins could be used to prevent  
 certain diseases and/or disease states. Thus, the ability to modulate the expression of  
 specific proteins in an organism would provide significant benefits.

Numerous methods have been developed to date for introducing foreign genes,  
 30 either under the control of an inducible, constitutively active, or endogenous  
 promoter, into organisms. Of particular interest are the inducible promoters (see, M.  
 Gossen, et al., *Proc. Natl. Acad. Sci. USA.*, 89:5547 (1992); Y. Wang, et al., *Proc.*

Natl. Acad. Sci. USA, 91:8180 (1994), D. No., et al., Proc. Natl. Acad. Sci. USA, 93:3346 (1996); and V.M. Rivera, et al., Nature Med, 2:1028 (1996); in addition to additional examples disclosed elsewhere herein). In one example, the gene for erythropoietin (Epo) was transferred into mice and primates under the control of a  
5 small molecule inducer for expression (e.g., tetracycline or rapamycin) (see, D. Bohl, et al., Blood, 92:1512, (1998); K.G. Rendahl, et al., Nat. Biotech, 16:757, (1998); V.M. Rivera, et al., Proc. Natl. Acad. Sci. USA, 96:8657 (1999); and X.Ye et al., Science, 283:88 (1999). Although such systems enable efficient induction of the gene of interest in the organism upon addition of the inducing agent (i.e., tetracycline,  
10 rapamycin, etc.), the levels of expression tend to peak at 24 hours and trail off to background levels after 4 to 14 days. Thus, controlled transient expression is virtually impossible using these systems, though such control would be desirable.

A new alternative method of controlling gene expression levels of a protein from a transgene (i.e., includes stable and transient transformants) has recently been  
15 elucidated (V.M. Rivera., et al., Science, 287:826-830, (2000)). This method does not control gene expression at the level of the mRNA like the aforementioned systems. Rather, the system controls the level of protein in an active secreted form. In the absence of the inducing agent, the protein aggregates in the ER and is not secreted. However, addition of the inducing agent results in dis-aggregation of the protein and  
20 the subsequent secretion from the ER. Such a system affords low basal secretion, rapid, high level secretion in the presence of the inducing agent, and rapid cessation of secretion upon removal of the inducing agent. In fact, protein secretion reached a maximum level within 30 minutes of induction, and a rapid cessation of secretion within 1 hour of removing the inducing agent. The method is also applicable for  
25 controlling the level of production for membrane proteins.

Detailed methods are presented in V.M. Rivera., et al., Science, 287:826-830, (2000)), briefly:

Fusion protein constructs are created using polynucleotide sequences of the present invention with one or more copies (preferably at least 2, 3, 4, or more) of a  
30 conditional aggregation domain (CAD) a domain that interacts with itself in a ligand-reversible manner (i.e., in the presence of an inducing agent) using molecular biology methods known in the art and discussed elsewhere herein. The CAD domain may be

the mutant domain isolated from the human FKBP12 (Phe<sup>36</sup> to Met) protein (as disclosed in V.M. Rivera., et al., Science, 287:826-830, (2000), or alternatively other proteins having domains with similar ligand-reversible, self-aggregation properties. As a principle of design the fusion protein vector would contain a furin cleavage sequence operably linked between the polynucleotides of the present invention and the CAD domains. Such a cleavage site would enable the proteolytic cleavage of the CAD domains from the polypeptide of the present invention subsequent to secretion from the ER and upon entry into the trans-Golgi (J.B. Denault, et al., FEBS Lett., 379:113, (1996)). Alternatively, the skilled artisan would recognize that any proteolytic cleavage sequence could be substituted for the furin sequence provided the substituted sequence is cleavable either endogenously (e.g., the furin sequence) or exogenously (e.g., post secretion, post purification, post production, etc.). The preferred sequence of each feature of the fusion protein construct, from the 5' to 3' direction with each feature being operably linked to the other, would be a promoter, signal sequence, "X" number of (CAD)x domains, the furin sequence (or other proteolytic sequence), and the coding sequence of the polypeptide of the present invention. The artisan would appreciate that the promotor and signal sequence, independent from the other, could be either the endogenous promotor or signal sequence of a polypeptide of the present invention, or alternatively, could be a heterologous signal sequence and promotor.

The specific methods described herein for controlling protein secretion levels through controlled ER aggregation are not meant to be limiting are would be generally applicable to any of the polynucleotides and polypeptides of the present invention, including variants, homologues, orthologs, and fragments therein.

#### EXAMPLE 20 – ALTERATION OF PROTEIN GLYCOSYLATION SITES TO ENHANCE CHARACTERISTICS OF POLYPEPTIDES OF THE INVENTION

Many eukaryotic cell surface and proteins are post-translationally processed to incorporate N-linked and O-linked carbohydrates (Kornfeld and Kornfeld (1985) Annu. Rev. Biochem. 54:631-64; Rademacher et al., (1988) Annu. Rev. Biochem. 57:785-838). Protein glycosylation is thought to serve a variety of functions including: augmentation of protein folding, inhibition of protein aggregation,

regulation of intracellular trafficking to organelles, increasing resistance to proteolysis, modulation of protein antigenicity, and mediation of intercellular adhesion (Fieldler and Simons (1995) *Cell*, 81:309-312; Helenius (1994) *Mol. Biol. Of the Cell* 5:253-265; Olden et al., (1978) *Cell*, 13:461-473; Caton et al., (1982) *Cell*, 37:417-427; Alexamnder and Elder (1984), *Science*, 226:1328-1330; and Flack et al., (1994), *J. Biol. Chem.*, 269:14015-14020). In higher organisms, the nature and extent of glycosylation can markedly affect the circulating half-life and bio-availability of proteins by mechanisms involving receptor mediated uptake and clearance (Ashwell and Morrell, (1974), *Adv. Enzymol.*, 41:99-128; Ashwell and Harford (1982), *Ann. Rev. Biochem.*, 51:531-54). Receptor systems have been identified that are thought to play a major role in the clearance of serum proteins through recognition of various carbohydrate structures on the glycoproteins (Stockert (1995), *Physiol. Rev.*, 75:591-609; Kery et al., (1992), *Arch. Biochem. Biophys.*, 298:49-55). Thus, production strategies resulting in incomplete attachment of terminal sialic acid residues might provide a means of shortening the bioavailability and half-life of glycoproteins. Conversely, expression strategies resulting in saturation of terminal sialic acid attachment sites might lengthen protein bioavailability and half-life.

In the development of recombinant glycoproteins for use as pharmaceutical products, for example, it has been speculated that the pharmacodynamics of recombinant proteins can be modulated by the addition or deletion of glycosylation sites from a glycoproteins primary structure (Berman and Lasky (1985a) *Trends in Biotechnol.*, 3:51-53). However, studies have reported that the deletion of N-linked glycosylation sites often impairs intracellular transport and results in the intracellular accumulation of glycosylation site variants (Machamer and Rose (1988), *J. Biol Chem.*, 263:5955-5960; Gallagher et al., (1992), *J. Virology.*, 66:7136-7145; Collier et al., (1993), *Biochem.*, 32:7818-7823; Claffey et al., (1995) *Biochemica et Biophysica Acta*, 1246:1-9; Dube et al., (1988), *J. Biol. Chem.* 263:17516-17521). While glycosylation site variants of proteins can be expressed intracellularly, it has proved difficult to recover useful quantities from growth conditioned cell culture medium.

Moreover, it is unclear to what extent a glycosylation site in one species will be recognized by another species glycosylation machinery. Due to the importance of

glycosylation in protein metabolism, particularly the secretion and/or expression of the protein, whether a glycosylation signal is recognized may profoundly determine a proteins ability to be expressed, either endogenously or recombinately, in another organism (i.e., expressing a human protein in E.coli, yeast, or viral organisms; or an  
5 E.coli, yeast, or viral protein in human, etc.). Thus, it may be desirable to add, delete, or modify a glycosylation site, and possibly add a glycosylation site of one species to a protein of another species to improve the proteins functional, bioprocess purification, and/or structural characteristics (e.g., a polypeptide of the present invention).

10 A number of methods may be employed to identify the location of glycosylation sites within a protein. One preferred method is to run the translated protein sequence through the PROSITE computer program (Swiss Institute of Bioinformatics). Once identified, the sites could be systematically deleted, or impaired, at the level of the DNA using mutagenesis methodology known in the art  
15 and available to the skilled artisan, Preferably using PCR-directed mutagenesis (See Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982)). Similarly, glycosylation sites could be added, or modified at the level of the DNA using similar methods, preferably PCR methods (See, Maniatis, supra). The results of modifying the glycosylation sites for a particular protein (e.g.,  
20 solubility, secretion potential, activity, aggregation, proteolytic resistance, etc.) could then be analyzed using methods know in the art.

The skilled artisan would acknowledge the existence of other computer algorithms capable of predicting the location of glycosylation sites within a protein. For example, the Motif computer program (Genetics Computer Group suite of  
25 programs) provides this function, as well.

#### EXAMPLE 21 – METHOD OF ENHANCING THE BIOLOGICAL ACTIVITY/FUNCTIONAL CHARACTERISTICS OF INVENTION THROUGH MOLECULAR EVOLUTION

30 Although many of the most biologically active proteins known are highly effective for their specified function in an organism, they often possess characteristics that make them undesirable for transgenic, therapeutic, and/or industrial applications.

Among these traits, a short physiological half-life is the most prominent problem, and is present either at the level of the protein, or the level of the proteins mRNA. The ability to extend the half-life, for example, would be particularly important for a proteins use in gene therapy, transgenic animal production, the bioprocess production and purification of the protein, and use of the protein as a chemical modulator among others. Therefore, there is a need to identify novel variants of isolated proteins possessing characteristics which enhance their application as a therapeutic for treating diseases of animal origin, in addition to the proteins applicability to common industrial and pharmaceutical applications.

Thus, one aspect of the present invention relates to the ability to enhance specific characteristics of invention through directed molecular evolution. Such an enhancement may, in a non-limiting example, benefit the inventions utility as an essential component in a kit, the inventions physical attributes such as its solubility, structure, or codon optimization, the inventions specific biological activity, including any associated enzymatic activity, the proteins enzyme kinetics, the proteins  $K_i$ ,  $K_{cat}$ ,  $K_m$ ,  $V_{max}$ ,  $K_d$ , protein-protein activity, protein-DNA binding activity, antagonist/inhibitory activity (including direct or indirect interaction), agonist activity (including direct or indirect interaction), the proteins antigenicity (e.g., where it would be desirable to either increase or decrease the antigenic potential of the protein), the immunogenicity of the protein, the ability of the protein to form dimers, trimers, or multimers with either itself or other proteins, the antigenic efficacy of the invention, including its subsequent use a preventative treatment for disease or disease states, or as an effector for targeting diseased genes. Moreover, the ability to enhance specific characteristics of a protein may also be applicable to changing the characterized activity of an enzyme to an activity completely unrelated to its initially characterized activity. Other desirable enhancements of the invention would be specific to each individual protein, and would thus be well known in the art and contemplated by the present invention.

For example, an engineered phosphatase may be constitutively active. Alternatively, an engineered phosphatase may be constitutively active in the absence of ligand binding. In yet another example, an engineered phosphatase may be capable of being activated with less than all of the regulatory factors and/or conditions

typically required for phosphatase activation (e.g., ligand binding, phosphorylation, conformational changes, etc.). Alternatively, an engineered phosphatase may have altered substrate specificity. Such phosphatases would be useful in screens to identify phosphatase modulators, among other uses described herein.

5 Directed evolution is comprised of several steps. The first step is to establish a library of variants for the gene or protein of interest. The most important step is to then select for those variants that entail the activity you wish to identify. The design of the screen is essential since your screen should be selective enough to eliminate non-useful variants, but not so stringent as to eliminate all variants. The last step is  
10 then to repeat the above steps using the best variant from the previous screen. Each successive cycle, can then be tailored as necessary, such as increasing the stringency of the screen, for example.

Over the years, there have been a number of methods developed to introduce mutations into macromolecules. Some of these methods include, random mutagenesis,  
15 “error-prone” PCR, chemical mutagenesis, site-directed mutagenesis, and other methods well known in the art (for a comprehensive listing of current mutagenesis methods, see Maniatis, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring, NY (1982)). Typically, such methods have been used, for example, as tools for identifying the core functional region(s) of a protein or the  
20 function of specific domains of a protein (if a multi-domain protein). However, such methods have more recently been applied to the identification of macromolecule variants with specific or enhanced characteristics.

Random mutagenesis has been the most widely recognized method to date. Typically, this has been carried out either through the use of “error-prone” PCR (as  
25 described in Moore, J., et al, *Nature Biotechnology* 14:458, (1996), or through the application of randomized synthetic oligonucleotides corresponding to specific regions of interest (as described by Derbyshire, K.M. et al, *Gene*, 46:145-152, (1986), and Hill, DE, et al, *Methods Enzymol.*, 55:559-568, (1987). Both approaches have limits to the level of mutagenesis that can be obtained. However, either approach  
30 enables the investigator to effectively control the rate of mutagenesis. This is particularly important considering the fact that mutations beneficial to the activity of

the enzyme are fairly rare. In fact, using too high a level of mutagenesis may counter or inhibit the desired benefit of a useful mutation.

While both of the aforementioned methods are effective for creating randomized pools of macromolecule variants, a third method, termed “DNA Shuffling”, or “sexual PCR” (WPC, Stemmer, PNAS, 91:10747, (1994)) has recently been elucidated. DNA shuffling has also been referred to as “directed molecular evolution”, “exon-shuffling”, “directed enzyme evolution”, “in vitro evolution”, and “artificial evolution”. Such reference terms are known in the art and are encompassed by the invention. This new, preferred, method apparently overcomes the limitations of the previous methods in that it not only propagates positive traits, but simultaneously eliminates negative traits in the resulting progeny.

DNA shuffling accomplishes this task by combining the principal of in vitro recombination, along with the method of “error-prone” PCR. In effect, you begin with a randomly digested pool of small fragments of your gene, created by Dnase I digestion, and then introduce said random fragments into an “error-prone” PCR assembly reaction. During the PCR reaction, the randomly sized DNA fragments not only hybridize to their cognate strand, but also may hybridize to other DNA fragments corresponding to different regions of the polynucleotide of interest – regions not typically accessible via hybridization of the entire polynucleotide. Moreover, since the PCR assembly reaction utilizes “error-prone” PCR reaction conditions, random mutations are introduced during the DNA synthesis step of the PCR reaction for all of the fragments -further diversifying the potential hybridization sites during the annealing step of the reaction.

A variety of reaction conditions could be utilized to carry-out the DNA shuffling reaction. However, specific reaction conditions for DNA shuffling are provided, for example, in PNAS, 91:10747, (1994). Briefly:

Prepare the DNA substrate to be subjected to the DNA shuffling reaction. Preparation may be in the form of simply purifying the DNA from contaminating cellular material, chemicals, buffers, oligonucleotide primers, deoxynucleotides, RNAs, etc., and may entail the use of DNA purification kits as those provided by Qiagen, Inc., or by the Promega, Corp., for example.

Once the DNA substrate has been purified, it would be subjected to Dnase I digestion. About 2-4ug of the DNA substrate(s) would be digested with .0015 units of Dnase I (Sigma) per ul in 100ul of 50mM Tris-HCL, pH 7.4/1mM MgCl<sub>2</sub> for 10-20 min. at room temperature. The resulting fragments of 10-50bp could then be purified  
5 by running them through a 2% low-melting point agarose gel by electrophoresis onto DE81 ion-exchange paper (Whatmann) or could be purified using Microcon concentrators (Amicon) of the appropriate molecular weight cutoff, or could use oligonucleotide purification columns (Qiagen), in addition to other methods known in the art. If using DE81 ion-exchange paper, the 10-50bp fragments could be eluted  
10 from said paper using 1M NaCl, followed by ethanol precipitation.

The resulting purified fragments would then be subjected to a PCR assembly reaction by re-suspension in a PCR mixture containing: 2mM of each dNTP, 2.2mM MgCl<sub>2</sub>, 50 mM KCl, 10mM Tris•HCL, pH 9.0, and 0.1% Triton X-100, at a final fragment concentration of 10-30ng/ul. No primers are added at this point. *Taq* DNA  
15 polymerase (Promega) would be used at 2.5 units per 100ul of reaction mixture. A PCR program of 94 C for 60s; 94 C for 30s, 50-55 C for 30s, and 72 C for 30s using 30-45 cycles, followed by 72 C for 5min using an MJ Research (Cambridge, MA) PTC-150 thermocycler. After the assembly reaction is completed, a 1:40 dilution of the resulting primerless product would then be introduced into a PCR mixture (using  
20 the same buffer mixture used for the assembly reaction) containing 0.8um of each primer and subjecting this mixture to 15 cycles of PCR (using 94 C for 30s, 50 C for 30s, and 72 C for 30s). The referred primers would be primers corresponding to the nucleic acid sequences of the polynucleotide(s) utilized in the shuffling reaction. Said primers could consist of modified nucleic acid base pairs using methods known in the  
25 art and referred to else where herein, or could contain additional sequences (i.e., for adding restriction sites, mutating specific base-pairs, etc.).

The resulting shuffled, assembled, and amplified product can be purified using methods well known in the art (e.g., Qiagen PCR purification kits) and then subsequently cloned using appropriate restriction enzymes.

30 Although a number of variations of DNA shuffling have been published to date, such variations would be obvious to the skilled artisan and are encompassed by the invention. The DNA shuffling method can also be tailored to the desired level of

mutagenesis using the methods described by Zhao, et al. (Nucl Acid Res., 25(6):1307-1308, (1997).

As described above, once the randomized pool has been created, it can then be subjected to a specific screen to identify the variant possessing the desired characteristic(s). Once the variant has been identified, DNA corresponding to the variant could then be used as the DNA substrate for initiating another round of DNA shuffling. This cycle of shuffling, selecting the optimized variant of interest, and then re-shuffling, can be repeated until the ultimate variant is obtained. Examples of model screens applied to identify variants created using DNA shuffling technology may be found in the following publications: J. C., Moore, et al., J. Mol. Biol., 272:336-347, (1997), F.R., Cross, et al., Mol. Cell. Biol., 18:2923-2931, (1998), and A. Cramer., et al., Nat. Biotech., 15:436-438, (1997).

DNA shuffling has several advantages. First, it makes use of beneficial mutations. When combined with screening, DNA shuffling allows the discovery of the best mutational combinations and does not assume that the best combination contains all the mutations in a population. Secondly, recombination occurs simultaneously with point mutagenesis. An effect of forcing DNA polymerase to synthesize full-length genes from the small fragment DNA pool is a background mutagenesis rate. In combination with a stringent selection method, enzymatic activity has been evolved up to 16000 fold increase over the wild-type form of the enzyme. In essence, the background mutagenesis yielded the genetic variability on which recombination acted to enhance the activity.

A third feature of recombination is that it can be used to remove deleterious mutations. As discussed above, during the process of the randomization, for every one beneficial mutation, there may be at least one or more neutral or inhibitory mutations. Such mutations can be removed by including in the assembly reaction an excess of the wild-type random-size fragments, in addition to the random-size fragments of the selected mutant from the previous selection. During the next selection, some of the most active variants of the polynucleotide/polypeptide/enzyme, should have lost the inhibitory mutations.

Finally, recombination enables parallel processing. This represents a significant advantage since there are likely multiple characteristics that would make a

protein more desirable (e.g. solubility, activity, etc.). Since it is increasingly difficult to screen for more than one desirable trait at a time, other methods of molecular evolution tend to be inhibitory. However, using recombination, it would be possible to combine the randomized fragments of the best representative variants for the various traits, and then select for multiple properties at once.

DNA shuffling can also be applied to the polynucleotides and polypeptides of the present invention to decrease their immunogenicity in a specified host. For example, a particular variant of the present invention may be created and isolated using DNA shuffling technology. Such a variant may have all of the desired characteristics, though may be highly immunogenic in a host due to its novel intrinsic structure. Specifically, the desired characteristic may cause the polypeptide to have a non-native structure which could no longer be recognized as a "self" molecule, but rather as a "foreign", and thus activate a host immune response directed against the novel variant. Such a limitation can be overcome, for example, by including a copy of the gene sequence for a xenobiotic ortholog of the native protein in with the gene sequence of the novel variant gene in one or more cycles of DNA shuffling. The molar ratio of the ortholog and novel variant DNAs could be varied accordingly. Ideally, the resulting hybrid variant identified would contain at least some of the coding sequence which enabled the xenobiotic protein to evade the host immune system, and additionally, the coding sequence of the original novel variant that provided the desired characteristics.

Likewise, the invention encompasses the application of DNA shuffling technology to the evolution of polynucleotides and polypeptides of the invention, wherein one or more cycles of DNA shuffling include, in addition to the gene template DNA, oligonucleotides coding for known allelic sequences, optimized codon sequences, known variant sequences, known polynucleotide polymorphism sequences, known ortholog sequences, known homologue sequences, additional homologous sequences, additional non-homologous sequences, sequences from another species, and any number and combination of the above.

In addition to the described methods above, there are a number of related methods that may also be applicable, or desirable in certain cases. Representative among these are the methods discussed in PCT applications WO 98/31700, and WO

98/32845, which are hereby incorporated by reference. Furthermore, related methods can also be applied to the polynucleotide sequences of the present invention in order to evolve invention for creating ideal variants for use in gene therapy, protein engineering, evolution of whole cells containing the variant, or in the evolution of entire enzyme pathways containing polynucleotides of the invention as described in PCT applications WO 98/13485, WO 98/13487, WO 98/27230, WO 98/31837, and Cramer, A., et al., Nat. Biotech., 15:436-438, (1997), respectively.

Additional methods of applying "DNA Shuffling" technology to the polynucleotides and polypeptides of the present invention, including their proposed applications, may be found in US Patent No. 5,605,793; PCT Application No. WO 95/22625; PCT Application No. WO 97/20078; PCT Application No. WO 97/35966; and PCT Application No. WO 98/42832; PCT Application No. WO 00/09727 specifically provides methods for applying DNA shuffling to the identification of herbicide selective crops which could be applied to the polynucleotides and polypeptides of the present invention; additionally, PCT Application No. WO 00/12680 provides methods and compositions for generating, modifying, adapting, and optimizing polynucleotide sequences that confer detectable phenotypic properties on plant species; each of the above are hereby incorporated in their entirety herein for all purposes.

#### EXAMPLE 22 – METHOD OF DETERMINING ALTERATIONS IN A GENE CORRESPONDING TO A POLYNUCLEOTIDE

RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:1. Suggested PCR conditions consist of 35 cycles at 95 degrees C for 30 seconds; 60-120 seconds at 52-58 degrees C; and 60-120 seconds at 70 degrees C, using buffer solutions described in Sidransky et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR

products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products are cloned into T-tailed vectors as described in Holton et al., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to the methods described herein are nick-translated with digoxigenin deoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

#### EXAMPLE 23 – METHOD OF DETECTING ABNORMAL LEVELS OF A POLYPEPTIDE IN A BIOLOGICAL SAMPLE

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described elsewhere  
5 herein. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled  
10 water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

15 Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale).  
20 Interpolate the concentration of the polypeptide in the sample using the standard curve.

#### EXAMPLE 24 – FORMULATION

The invention also provides methods of treatment and/or prevention diseases,  
25 disorders, and/or conditions (such as, for example, any one or more of the diseases or disorders disclosed herein) by administration to a subject of an effective amount of a Therapeutic. By therapeutic is meant a polynucleotides or polypeptides of the invention (including fragments and variants), agonists or antagonists thereof, and/or antibodies thereto, in combination with a pharmaceutically acceptable carrier type  
30 (e.g., a sterile carrier).

The Therapeutic will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual

patient (especially the side effects of treatment with the Therapeutic alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

5           As a general proposition, the total pharmaceutically effective amount of the Therapeutic administered parenterally per dose will be in the range of about 1ug/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for  
10 the hormone. If given continuously, the Therapeutic is typically administered at a dose rate of about 1 ug/kg/hour to about 50 ug/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to  
15 vary depending on the desired effect.

          Therapeutics can be administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating  
20 material or formulation auxiliary of any. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

          Therapeutics of the invention are also suitably administered by sustained-release systems. Suitable examples of sustained-release Therapeutics are administered  
25 orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of  
30 administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

Therapeutics of the invention may also be suitably administered by sustained-release systems. Suitable examples of sustained-release Therapeutics include suitable polymeric materials (such as, for example, semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules), suitable hydrophobic materials  
 5 (for example as an emulsion in an acceptable oil) or ion exchange resins, and sparingly soluble derivatives (such as, for example, a sparingly soluble salt).

Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman et al., Biopolymers 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (Langer et al.,  
 10 J. Biomed. Mater. Res. 15:167-277 (1981), and Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (Langer et al., Id.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988).

Sustained-release Therapeutics also include liposomally entrapped Therapeutics of the invention (see, generally, Langer, Science 249:1527-1533 (1990);  
 15 Treat et al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 317 -327 and 353-365 (1989)). Liposomes containing the Therapeutic are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. (USA) 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci.(USA) 77:4030-4034 (1980); EP 52,322; EP 36,676; EP  
 20 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal Therapeutic.

25 In yet an additional embodiment, the Therapeutics of the invention are delivered by way of a pump (see Langer, supra; Sefton, CRC Crit. Ref. Biomed. Eng. 14:201 (1987); Buchwald et al., Surgery 88:507 (1980); Saudek et al., N. Engl. J. Med. 321:574 (1989)).

Other controlled release systems are discussed in the review by Langer  
 30 (Science 249:1527-1533 (1990)).

For parenteral administration, in one embodiment, the Therapeutic is formulated generally by mixing it at the desired degree of purity, in a unit dosage

injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other  
5 compounds that are known to be deleterious to the Therapeutic.

Generally, the formulations are prepared by contacting the Therapeutic uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood  
10 of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to  
15 recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as  
20 polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

25 The Therapeutic will typically be formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

Any pharmaceutical used for therapeutic administration can be sterile. Sterility  
30 is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutics generally are placed into a container having a

sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

Therapeutics ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized  
5 formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous Therapeutic solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized Therapeutic using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or  
10 more containers filled with one or more of the ingredients of the Therapeutics of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the Therapeutics may be employed in  
15 conjunction with other therapeutic compounds.

The Therapeutics of the invention may be administered alone or in combination with adjuvants. Adjuvants that may be administered with the Therapeutics of the invention include, but are not limited to, alum, alum plus deoxycholate (ImmunoAg), MTP-PE (Biocine Corp.), QS21 (Genentech, Inc.), BCG,  
20 and MPL. In a specific embodiment, Therapeutics of the invention are administered in combination with alum. In another specific embodiment, Therapeutics of the invention are administered in combination with QS-21. Further adjuvants that may be administered with the Therapeutics of the invention include, but are not limited to, Monophosphoryl lipid immunomodulator, AdjuVax 100a, QS-21, QS-18, CRL1005,  
25 Aluminum salts, MF-59, and Virosomal adjuvant technology. Vaccines that may be administered with the Therapeutics of the invention include, but are not limited to, vaccines directed toward protection against MMR (measles, mumps, rubella), polio, varicella, tetanus/diphtheria, hepatitis A, hepatitis B, haemophilus influenzae B, whooping cough, pneumonia, influenza, Lyme's Disease, rotavirus, cholera, yellow  
30 fever, Japanese encephalitis, poliomyelitis, rabies, typhoid fever, and pertussis. Combinations may be administered either concomitantly, e.g., as an admixture, separately but simultaneously or concurrently; or sequentially. This includes

presentations in which the combined agents are administered together as a therapeutic mixture, and also procedures in which the combined agents are administered separately but simultaneously, e.g., as through separate intravenous lines into the same individual. Administration "in combination" further includes the separate  
5 administration of one of the compounds or agents given first, followed by the second.

The Therapeutics of the invention may be administered alone or in combination with other therapeutic agents. Therapeutic agents that may be administered in combination with the Therapeutics of the invention, include but not limited to, other members of the TNF family, chemotherapeutic agents, antibiotics,  
10 steroidal and non-steroidal anti-inflammatories, conventional immunotherapeutic agents, cytokines and/or growth factors. Combinations may be administered either concomitantly, e.g., as an admixture, separately but simultaneously or concurrently; or sequentially. This includes presentations in which the combined agents are administered together as a therapeutic mixture, and also procedures in which the  
15 combined agents are administered separately but simultaneously, e.g., as through separate intravenous lines into the same individual. Administration "in combination" further includes the separate administration of one of the compounds or agents given first, followed by the second.

In one embodiment, the Therapeutics of the invention are administered in combination with members of the TNF family. TNF, TNF-related or TNF-like  
20 molecules that may be administered with the Therapeutics of the invention include, but are not limited to, soluble forms of TNF-alpha, lymphotoxin-alpha (LT-alpha, also known as TNF-beta), LT-beta (found in complex heterotrimer LT-alpha2-beta), OPGL, FasL, CD27L, CD30L, CD40L, 4-1BBL, DcR3, OX40L, TNF-gamma  
25 (International Publication No. WO 96/14328), AIM-I (International Publication No. WO 97/33899), endokine-alpha (International Publication No. WO 98/07880), TR6 (International Publication No. WO 98/30694), OPG, and neutrokin-alpha (International Publication No. WO 98/18921, OX40, and nerve growth factor (NGF), and soluble forms of Fas, CD30, CD27, CD40 and 4-IBB, TR2 (International  
30 Publication No. WO 96/34095), DR3 (International Publication No. WO 97/33904), DR4 (International Publication No. WO 98/32856), TR5 (International Publication No. WO 98/30693), TR6 (International Publication No. WO 98/30694), TR7

(International Publication No. WO 98/41629), TRANK, TR9 (International Publication No. WO 98/56892), TR10 (International Publication No. WO 98/54202), 312C2 (International Publication No. WO 98/06842), and TR12, and soluble forms CD154, CD70, and CD153.

5 In certain embodiments, Therapeutics of the invention are administered in combination with antiretroviral agents, nucleoside reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors, and/or protease inhibitors. Nucleoside reverse transcriptase inhibitors that may be administered in combination with the Therapeutics of the invention, include, but are not limited to, RETROVIR  
 10 (zidovudine/AZT), VIDEX (didanosine/ddI), HIVID (zalcitabine/ddC), ZERIT (stavudine/d4T), EPIVIR (lamivudine/3TC), and COMBIVIR (zidovudine/lamivudine). Non-nucleoside reverse transcriptase inhibitors that may be administered in combination with the Therapeutics of the invention, include, but are not limited to, VIRAMUNE (nevirapine), RESCRIPTOR (delavirdine), and  
 15 SUSTIVA (efavirenz). Protease inhibitors that may be administered in combination with the Therapeutics of the invention, include, but are not limited to, CRIXIVAN (indinavir), NORVIR (ritonavir), INVIRASE (saquinavir), and VIRACEPT (nelfinavir). In a specific embodiment, antiretroviral agents, nucleoside reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors, and/or  
 20 protease inhibitors may be used in any combination with Therapeutics of the invention to treat AIDS and/or to prevent or treat HIV infection.

In other embodiments, Therapeutics of the invention may be administered in combination with anti-opportunistic infection agents. Anti-opportunistic agents that may be administered in combination with the Therapeutics of the invention, include,  
 25 but are not limited to, TRIMETHOPRIM-SULFAMETHOXAZOLE, DAPSONE, PENTAMIDINE, ATOVAQUONE, ISONIAZID, RIFAMPIN, PYRAZINAMIDE, ETHAMBUTOL, RIFABUTIN, CLARITHROMYCIN, AZITHROMYCIN, GANCICLOVIR, FOSCARNET, CIDOFOVIR, FLUCONAZOLE, ITRACONAZOLE, KETOCONAZOLE, ACYCLOVIR, FAMCICLOVIR,  
 30 PYRIMETHAMINE, LEUCOVORIN, NEUPOGEN (filgrastim/G-CSF), and LEUKINE (sargramostim/GM-CSF). In a specific embodiment, Therapeutics of the invention are used in any combination with TRIMETHOPRIM-

SULFAMETHOXAZOLE, DAPSONE, PENTAMIDINE, and/or ATOVAQUONE to prophylactically treat or prevent an opportunistic *Pneumocystis carinii* pneumonia infection. In another specific embodiment, Therapeutics of the invention are used in any combination with ISONIAZID, RIFAMPIN, PYRAZINAMIDE, and/or

5 ETHAMBUTOL to prophylactically treat or prevent an opportunistic *Mycobacterium avium* complex infection. In another specific embodiment, Therapeutics of the invention are used in any combination with RIFABUTIN, CLARITHROMYCIN, and/or AZITHROMYCIN to prophylactically treat or prevent an opportunistic *Mycobacterium tuberculosis* infection. In another specific embodiment, Therapeutics

10 of the invention are used in any combination with GANCICLOVIR, FOSCARNET, and/or CIDOFOVIR to prophylactically treat or prevent an opportunistic cytomegalovirus infection. In another specific embodiment, Therapeutics of the invention are used in any combination with FLUCONAZOLE, ITRACONAZOLE, and/or KETOCONAZOLE to prophylactically treat or prevent an opportunistic fungal

15 infection. In another specific embodiment, Therapeutics of the invention are used in any combination with ACYCLOVIR and/or FAMCICOLVIR to prophylactically treat or prevent an opportunistic herpes simplex virus type I and/or type II infection. In another specific embodiment, Therapeutics of the invention are used in any combination with PYRIMETHAMINE and/or LEUCOVORIN to prophylactically

20 treat or prevent an opportunistic *Toxoplasma gondii* infection. In another specific embodiment, Therapeutics of the invention are used in any combination with LEUCOVORIN and/or NEUPOGEN to prophylactically treat or prevent an opportunistic bacterial infection.

In a further embodiment, the Therapeutics of the invention are administered in

25 combination with an antiviral agent. Antiviral agents that may be administered with the Therapeutics of the invention include, but are not limited to, acyclovir, ribavirin, amantadine, and remantidine.

In a further embodiment, the Therapeutics of the invention are administered in combination with an antibiotic agent. Antibiotic agents that may be administered with

30 the Therapeutics of the invention include, but are not limited to, amoxicillin, beta-lactamases, aminoglycosides, beta-lactam (glycopeptide), beta-lactamases, Clindamycin, chloramphenicol, cephalosporins, ciprofloxacin, ciprofloxacin,

erythromycin, fluoroquinolones, macrolides, metronidazole, penicillins, quinolones, rifampin, streptomycin, sulfonamide, tetracyclines, trimethoprim, trimethoprim-sulfamthoxazole, and vancomycin.

5 Conventional nonspecific immunosuppressive agents, that may be administered in combination with the Therapeutics of the invention include, but are not limited to, steroids, cyclosporine, cyclosporine analogs, cyclophosphamide methylprednisone, prednisone, azathioprine, FK-506, 15-deoxyspergualin, and other immunosuppressive agents that act by suppressing the function of responding T cells.

10 In specific embodiments, Therapeutics of the invention are administered in combination with immunosuppressants. Immunosuppressants preparations that may be administered with the Therapeutics of the invention include, but are not limited to, ORTHOCLONE (OKT3), SANDIMMUNE/NEORAL/SANGDYA (cyclosporin), PROGRAF (tacrolimus), CELLCEPT (mycophenolate), Azathioprine, glucocorticosteroids, and RAPAMUNE (sirolimus). In a specific embodiment, 15 immunosuppressants may be used to prevent rejection of organ or bone marrow transplantation.

In an additional embodiment, Therapeutics of the invention are administered alone or in combination with one or more intravenous immune globulin preparations. Intravenous immune globulin preparations that may be administered with the 20 Therapeutics of the invention include, but not limited to, GAMMAR, IVEEGAM, SANDOGLOBULIN, GAMMAGARD S/D, and GAMIMUNE. In a specific embodiment, Therapeutics of the invention are administered in combination with intravenous immune globulin preparations in transplantation therapy (e.g., bone marrow transplant).

25 In an additional embodiment, the Therapeutics of the invention are administered alone or in combination with an anti-inflammatory agent. Anti-inflammatory agents that may be administered with the Therapeutics of the invention include, but are not limited to, glucocorticoids and the nonsteroidal anti-inflammatories, aminoarylcarboxylic acid derivatives, arylacetic acid derivatives, 30 arylbutyric acid derivatives, arylcarboxylic acids, arylpropionic acid derivatives, pyrazoles, pyrazolones, salicylic acid derivatives, thiazinecarboxamides, e-acetamidocaproic acid, S-adenosylmethionine, 3-amino-4-hydroxybutyric acid,

amixetrine, bendazac, benzydamine, bucolome, difenpiramide, ditazol, emorfazone, guaiazulene, nabumetone, nimesulide, orgotein, oxaceprol, paranyline, perisoxal, pifoxime, proquazone, proxazole, and tenidap.

In another embodiment, compositions of the invention are administered in combination with a chemotherapeutic agent. Chemotherapeutic agents that may be administered with the Therapeutics of the invention include, but are not limited to, antibiotic derivatives (e.g., doxorubicin, bleomycin, daunorubicin, and dactinomycin); antiestrogens (e.g., tamoxifen); antimetabolites (e.g., fluorouracil, 5-FU, methotrexate, floxuridine, interferon alpha-2b, glutamic acid, plicamycin, mercaptopurine, and 6-thioguanine); cytotoxic agents (e.g., carmustine, BCNU, lomustine, CCNU, cytosine arabinoside, cyclophosphamide, estramustine, hydroxyurea, procarbazine, mitomycin, busulfan, cis-platin, and vincristine sulfate); hormones (e.g., medroxyprogesterone, estramustine phosphate sodium, ethinyl estradiol, estradiol, megestrol acetate, methyltestosterone, diethylstilbestrol diphosphate, chlorotrianisene, and testolactone); nitrogen mustard derivatives (e.g., mephallen, chorambucil, mechlorethamine (nitrogen mustard) and thiotepa); steroids and combinations (e.g., bethamethasone sodium phosphate); and others (e.g., dicarbazine, asparaginase, mitotane, vincristine sulfate, vinblastine sulfate, and etoposide).

In a specific embodiment, formulations of the present invention may further comprise antagonists of P-glycoprotein (also referred to as the multidrug resistance protein, or PGP), including antagonists of its encoding polynucleotides (e.g., antisense oligonucleotides, ribozymes, zinc-finger proteins, etc.). P-glycoprotein is well known for decreasing the efficacy of various drug administrations due to its ability to export intracellular levels of absorbed drug to the cell exterior. While this activity has been particularly pronounced in cancer cells in response to the administration of chemotherapy regimens, a variety of other cell types and the administration of other drug classes have been noted (e.g., T-cells and anti-HIV drugs). In fact, certain mutations in the PGP gene significantly reduces PGP function, making it less able to force drugs out of cells. People who have two versions of the mutated gene--one inherited from each parent--have more than four times less PGP than those with two normal versions of the gene. People may also have one normal gene and one mutated

one. Certain ethnic populations have increased incidence of such PGP mutations. Among individuals from Ghana, Kenya, the Sudan, as well as African Americans, frequency of the normal gene ranged from 73% to 84%. In contrast, the frequency was 34% to 59% among British whites, Portuguese, Southwest Asian, Chinese, Filipino and Saudi populations. As a result, certain ethnic populations may require increased administration of PGP antagonist in the formulation of the present invention to arrive at the an efficacious dose of the therapeutic (e.g., those from African descent). Conversely, certain ethnic populations, particularly those having increased frequency of the mutated PGP (e.g., of Caucasian descent, or non-African descent) may require less pharmaceutical compositions in the formulation due to an effective increase in efficacy of such compositions as a result of the increased effective absorption (e.g., less PGP activity) of said composition.

Moreover, in another specific embodiment, formulations of the present invention may further comprise antagonists of OATP2 (also referred to as the multiresistance protein, or MRP2), including antagonists of its encoding polynucleotides (e.g., antisense oligonucleotides, ribozymes, zinc-finger proteins, etc.). The invention also further comprises any additional antagonists known to inhibit proteins thought to be attributable to a multidrug resistant phenotype in proliferating cells.

In a specific embodiment, Therapeutics of the invention are administered in combination with CHOP (cyclophosphamide, doxorubicin, vincristine, and prednisone) or any combination of the components of CHOP. In another embodiment, Therapeutics of the invention are administered in combination with Rituximab. In a further embodiment, Therapeutics of the invention are administered with Rituxmab and CHOP, or Rituxmab and any combination of the components of CHOP.

In an additional embodiment, the Therapeutics of the invention are administered in combination with cytokines. Cytokines that may be administered with the Therapeutics of the invention include, but are not limited to, IL2, IL3, IL4, IL5, IL6, IL7, IL10, IL12, IL13, IL15, anti-CD40, CD40L, IFN-gamma and TNF-alpha. In another embodiment, Therapeutics of the invention may be administered with any interleukin, including, but not limited to, IL-1alpha, IL-1beta, IL-2, IL-3, IL-4, IL-5,

IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-19, IL-20, and IL-21.

In an additional embodiment, the Therapeutics of the invention are administered in combination with angiogenic proteins. Angiogenic proteins that may be administered with the Therapeutics of the invention include, but are not limited to, Glioma Derived Growth Factor (GDGF), as disclosed in European Patent Number EP-399816; Platelet Derived Growth Factor-A (PDGF-A), as disclosed in European Patent Number EP-682110; Platelet Derived Growth Factor-B (PDGF-B), as disclosed in European Patent Number EP-282317; Placental Growth Factor (PlGF), as disclosed in International Publication Number WO 92/06194; Placental Growth Factor-2 (PlGF-2), as disclosed in Hauser et al., Growth Factors, 4:259-268 (1993); Vascular Endothelial Growth Factor (VEGF), as disclosed in International Publication Number WO 90/13649; Vascular Endothelial Growth Factor-A (VEGF-A), as disclosed in European Patent Number EP-506477; Vascular Endothelial Growth Factor-2 (VEGF-2), as disclosed in International Publication Number WO 96/39515; Vascular Endothelial Growth Factor B (VEGF-3); Vascular Endothelial Growth Factor B-186 (VEGF-B186), as disclosed in International Publication Number WO 96/26736; Vascular Endothelial Growth Factor-D (VEGF-D), as disclosed in International Publication Number WO 98/02543; Vascular Endothelial Growth Factor-D (VEGF-D), as disclosed in International Publication Number WO 98/07832; and Vascular Endothelial Growth Factor-E (VEGF-E), as disclosed in German Patent Number DE19639601. The above mentioned references are incorporated herein by reference herein.

In an additional embodiment, the Therapeutics of the invention are administered in combination with hematopoietic growth factors. Hematopoietic growth factors that may be administered with the Therapeutics of the invention include, but are not limited to, LEUKINE (SARGRAMOSTIM) and NEUPOGEN (FILGRASTIM).

In an additional embodiment, the Therapeutics of the invention are administered in combination with Fibroblast Growth Factors. Fibroblast Growth Factors that may be administered with the Therapeutics of the invention include, but

are not limited to, FGF-1, FGF-2, FGF-3, FGF-4, FGF-5, FGF-6, FGF-7, FGF-8, FGF-9, FGF-10, FGF-11, FGF-12, FGF-13, FGF-14, and FGF-15.

In a specific embodiment, formulations of the present invention may further comprise antagonists of P-glycoprotein (also referred to as the multiresistance protein, or PGP), including antagonists of its encoding polynucleotides (e.g., antisense oligonucleotides, ribozymes, zinc-finger proteins, etc.). P-glycoprotein is well known for decreasing the efficacy of various drug administrations due to its ability to export intracellular levels of absorbed drug to the cell exterior. While this activity has been particularly pronounced in cancer cells in response to the administration of chemotherapy regimens, a variety of other cell types and the administration of other drug classes have been noted (e.g., T-cells and anti-HIV drugs). In fact, certain mutations in the PGP gene significantly reduces PGP function, making it less able to force drugs out of cells. People who have two versions of the mutated gene--one inherited from each parent--have more than four times less PGP than those with two normal versions of the gene. People may also have one normal gene and one mutated one. Certain ethnic populations have increased incidence of such PGP mutations. Among individuals from Ghana, Kenya, the Sudan, as well as African Americans, frequency of the normal gene ranged from 73% to 84%. In contrast, the frequency was 34% to 59% among British whites, Portuguese, Southwest Asian, Chinese, Filipino and Saudi populations. As a result, certain ethnic populations may require increased administration of PGP antagonist in the formulation of the present invention to arrive at the an efficacious dose of the therapeutic (e.g., those from African descent). Conversely, certain ethnic populations, particularly those having increased frequency of the mutated PGP (e.g., of Caucasian descent, or non-African descent) may require less pharmaceutical compositions in the formulation due to an effective increase in efficacy of such compositions as a result of the increased effective absorption (e.g., less PGP activity) of said composition.

Moreover, in another specific embodiment, formulations of the present invention may further comprise antagonists of OATP2 (also referred to as the multiresistance protein, or MRP2), including antagonists of its encoding polynucleotides (e.g., antisense oligonucleotides, ribozymes, zinc-finger proteins, etc.). The invention also further comprises any additional antagonists known to inhibit

In additional embodiments, the Therapeutics of the invention are administered in combination with other therapeutic or prophylactic regimens, such as, for example, radiation therapy.

The present invention relates to a method for treating an individual in need of an increased level of a polypeptide of the invention in the body comprising administering to such an individual a composition comprising a therapeutically effective amount of an agonist of the invention (including polypeptides of the invention). Moreover, it will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a Therapeutic comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided herein.

30           The present invention also relates to a method of treating an individual in need of a decreased level of a polypeptide of the invention in the body comprising administering to such an individual a composition comprising a therapeutically

effective amount of an antagonist of the invention (including polypeptides and antibodies of the invention).

In one example, antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer. For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided herein.

#### EXAMPLE 27 – METHOD OF TREATMENT USING GENE THERAPY-EX VIVO

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37 degree C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 12 using primers and having appropriate restriction sites and

initiation/stop codons, if necessary. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

#### EXAMPLE 28 – GENE THERAPY USING ENDOGENOUS GENES

#### CORRESPONDING TO POLYNUCLEOTIDES OF THE INVENTION

Another method of gene therapy according to the present invention involves operably associating the endogenous polynucleotide sequence of the invention with a

promoter via homologous recombination as described, for example, in U.S. Patent NO: 5,641,670, issued June 24, 1997; International Publication NO: WO 96/29411, published September 26, 1996; International Publication NO: WO 94/12650, published August 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA, 86:8932-8935  
5 (1989); and Zijlstra et al., Nature, 342:435-438 (1989). This method involves the activation of a gene which is present in the target cells, but which is not expressed in the cells, or is expressed at a lower level than desired.

Polynucleotide constructs are made which contain a promoter and targeting sequences, which are homologous to the 5' non-coding sequence of endogenous  
10 polynucleotide sequence, flanking the promoter. The targeting sequence will be sufficiently near the 5' end of the polynucleotide sequence so the promoter will be operably linked to the endogenous sequence upon homologous recombination. The promoter and the targeting sequences can be amplified using PCR. Preferably, the amplified promoter contains distinct restriction enzyme sites on the 5' and 3' ends.  
15 Preferably, the 3' end of the first targeting sequence contains the same restriction enzyme site as the 5' end of the amplified promoter and the 5' end of the second targeting sequence contains the same restriction site as the 3' end of the amplified promoter.

The amplified promoter and the amplified targeting sequences are digested  
20 with the appropriate restriction enzymes and subsequently treated with calf intestinal phosphatase. The digested promoter and digested targeting sequences are added together in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The construct is size fractionated on an agarose gel then purified by phenol extraction and ethanol  
25 precipitation.

In this Example, the polynucleotide constructs are administered as naked polynucleotides via electroporation. However, the polynucleotide constructs may also be administered with transfection-facilitating agents, such as liposomes, viral sequences, viral particles, precipitating agents, etc. Such methods of delivery are  
30 known in the art.

Once the cells are transfected, homologous recombination will take place which results in the promoter being operably linked to the endogenous polynucleotide

sequence. This results in the expression of polynucleotide corresponding to the polynucleotide in the cell. Expression may be detected by immunological staining, or any other method known in the art.

5 Fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in DMEM + 10% fetal calf serum. Exponentially growing or early stationary phase fibroblasts are trypsinized and rinsed from the plastic surface with nutrient medium. An aliquot of the cell suspension is removed for counting, and the remaining cells are subjected to centrifugation. The supernatant is aspirated and the pellet is resuspended in 5 ml of electroporation buffer (20 mM HEPES pH 7.3, 137 mM NaCl,  
10 5 mM KCl, 0.7 mM Na<sub>2</sub> HPO<sub>4</sub>, 6 mM dextrose). The cells are recentrifuged, the supernatant aspirated, and the cells resuspended in electroporation buffer containing 1 mg/ml acetylated bovine serum albumin. The final cell suspension contains approximately 3X10<sup>6</sup> cells/ml. Electroporation should be performed immediately following resuspension.

15 Plasmid DNA is prepared according to standard techniques. For example, to construct a plasmid for targeting to the locus corresponding to the polynucleotide of the invention, plasmid pUC18 (MBI Fermentas, Amherst, NY) is digested with HindIII. The CMV promoter is amplified by PCR with an XbaI site on the 5' end and a BamHI site on the 3' end. Two non-coding sequences are amplified via PCR: one  
20 non-coding sequence (fragment 1) is amplified with a HindIII site at the 5' end and an Xba site at the 3' end; the other non-coding sequence (fragment 2) is amplified with a BamHI site at the 5' end and a HindIII site at the 3' end. The CMV promoter and the fragments (1 and 2) are digested with the appropriate enzymes (CMV promoter - XbaI and BamHI; fragment 1 - XbaI; fragment 2 - BamHI) and ligated together. The  
25 resulting ligation product is digested with HindIII, and ligated with the HindIII-digested pUC18 plasmid.

Plasmid DNA is added to a sterile cuvette with a 0.4 cm electrode gap (Bio-Rad). The final DNA concentration is generally at least 120 µg/ml. 0.5 ml of the cell suspension (containing approximately 1.5X10<sup>6</sup> cells) is then added to the cuvette,  
30 and the cell suspension and DNA solutions are gently mixed. Electroporation is performed with a Gene-Pulser apparatus (Bio-Rad). Capacitance and voltage are set at 960 µF and 250-300 V, respectively. As voltage increases, cell survival decreases, but

the percentage of surviving cells that stably incorporate the introduced DNA into their genome increases dramatically. Given these parameters, a pulse time of approximately 14-20 mSec should be observed.

5 Electroporated cells are maintained at room temperature for approximately 5 min, and the contents of the cuvette are then gently removed with a sterile transfer pipette. The cells are added directly to 10 ml of prewarmed nutrient media (DMEM with 15% calf serum) in a 10 cm dish and incubated at 37 degree C. The following day, the media is aspirated and replaced with 10 ml of fresh media and incubated for a further 16-24 hours.

10 The engineered fibroblasts are then injected into the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads. The fibroblasts now produce the protein product. The fibroblasts can then be introduced into a patient as described above.

15 **EXAMPLE 29 – METHOD OF TREATMENT USING  
GENE THERAPY - IN VIVO**

Another aspect of the present invention is using in vivo gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA)  
20 sequences into an animal to increase or decrease the expression of the polypeptide. The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622,  
25 5705151, 5580859; Tabata et al., Cardiovasc. Res. 35(3):470-479 (1997); Chao et al., Pharmacol. Res. 35(6):517-522 (1997); Wolff, Neuromuscul. Disord. 7(5):314-318 (1997); Schwartz et al., Gene Ther. 3(5):405-411 (1996); Tsurumi et al., Circulation 94(12):3281-3290 (1996) (incorporated herein by reference).

30 The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The

polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Felgner P.L. et al. (1995) *Ann. NY Acad. Sci.* 772:126-139 and Abdallah B. et al. (1995) *Biol. Cell* 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated

cells, such as, for example, stem cells of blood or skin fibroblasts. In vivo muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle in vivo is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that

quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

### EXAMPLE 30 – TRANSGENIC ANIMALS

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (i.e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., *Appl. Microbiol. Biotechnol.* 40:691-698 (1994); Carver et al., *Biotechnology (NY)* 11:1263-1270 (1993); Wright et al., *Biotechnology (NY)* 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus mediated gene transfer into germ lines (Van der Putten et al., *Proc. Natl. Acad. Sci., USA* 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., *Cell* 56:313-321 (1989)); electroporation of cells or embryos (Lo, 1983, *Mol Cell. Biol.* 3:1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e.g., Ulmer et al., *Science* 259:1745 (1993); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (Lavitrano et al., *Cell* 57:717-723 (1989); etc. For a review of such techniques, see Gordon, "Transgenic Animals," *Intl. Rev. Cytol.* 115:171-229 (1989), which is incorporated by reference herein in its entirety.

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., Nature 380:64-66 (1996); Wilmut et al., Nature 385:810-813 (1997)).

The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e., mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89:6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., Science 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and reverse

transcriptase-PCR(RT-PCR).. Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying diseases, disorders, and/or conditions associated with aberrant expression, and in screening for compounds effective in ameliorating such diseases, disorders, and/or conditions.

#### EXAMPLE 31 – KNOCK-OUT ANIMALS

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (E.g., see Smithies et al., Nature 317:230-234 (1985); Thomas & Capecchi, Cell 51:503-512 (1987); Thompson et al., Cell 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention in vivo. In another embodiment, techniques known in the art are used to

generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e.g., see Thomas & Capecchi 1987 and Thompson 1989, supra). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors that will be apparent to those of skill in the art.

10           In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e.g., knockouts) are administered to a patient in vivo. Such cells may be obtained from the patient (i.e., animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e.g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered in vitro using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc. The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or intraperitoneally.

25           Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Patent No. 5,399,349; and

Mulligan & Wilson, U.S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying diseases, disorders, and/or conditions associated with aberrant expression, and in screening for compounds effective in ameliorating such diseases, disorders, and/or conditions.

#### EXAMPLE 32 – METHOD OF ISOLATING ANTIBODY FRAGMENTS DIRECTED AGAINST BMY\_HPP13 FROM A LIBRARY OF scFvs

Naturally occurring V-genes isolated from human PBLs are constructed into a library of antibody fragments which contain reactivities against BMY\_HPP13 to which the donor may or may not have been exposed (see e.g., U.S. Patent 5,885,793 incorporated herein by reference in its entirety).

Rescue of the Library. A library of scFvs is constructed from the RNA of human PBLs as described in PCT publication WO 92/01047. To rescue phage displaying antibody fragments, approximately 10<sup>9</sup> E. coli harboring the phagemid are used to inoculate 50 ml of 2xTY containing 1% glucose and 100 µg/ml of ampicillin (2xTY-AMP-GLU) and grown to an O.D. of 0.8 with shaking. Five ml of this culture is used to inoculate 50 ml of 2xTY-AMP-GLU, 2 x 10<sup>8</sup> TU of delta gene 3 helper (M13 delta gene III, see PCT publication WO 92/01047) are added and the culture incubated at 37°C for 45 minutes without shaking and then at 37°C for 45 minutes with shaking. The culture is centrifuged at 4000 r.p.m. for 10 min. and the pellet resuspended in 2 liters of 2xTY containing 100 µg/ml ampicillin and 50 ug/ml kanamycin and grown overnight. Phage are prepared as described in PCT publication WO 92/01047.

M13 delta gene III is prepared as follows: M13 delta gene III helper phage does not encode gene III protein, hence the phage(mid) displaying antibody fragments have a greater avidity of binding to antigen. Infectious M13 delta gene III particles are made by growing the helper phage in cells harboring a pUC19 derivative supplying the wild type gene III protein during phage morphogenesis. The culture is incubated for 1 hour at 37° C without shaking and then for a further hour at 37°C with shaking. Cells are spun down (IEC-Centra 8,400 r.p.m. for 10 min), resuspended in 300 ml 2xTY broth containing 100 µg ampicillin/ml and 25 µg kanamycin/ml (2xTY-AMP-KAN) and grown overnight, shaking at 37°C. Phage particles are purified and concentrated from the culture medium by two PEG-precipitations (Sambrook et al., 1990), resuspended in 2 ml PBS and passed through a 0.45 µm filter (Minisart NML; Sartorius) to give a final concentration of approximately 10<sup>13</sup> transducing units/ml (ampicillin-resistant clones).

Panning of the Library. Immunotubes (Nunc) are coated overnight in PBS with 4 ml of either 100 µg/ml or 10 µg/ml of a polypeptide of the present invention. Tubes are blocked with 2% Marvel-PBS for 2 hours at 37°C and then washed 3 times in PBS. Approximately 10<sup>13</sup> TU of phage is applied to the tube and incubated for 30 minutes at room temperature tumbling on an over and under turntable and then left to stand for another 1.5 hours. Tubes are washed 10 times with PBS 0.1% Tween-20 and 10 times with PBS. Phage are eluted by adding 1 ml of 100 mM triethylamine and rotating 15 minutes on an under and over turntable after which the solution is immediately neutralized with 0.5 ml of 1.0M Tris-HCl, pH 7.4. Phage are then used to infect 10 ml of mid-log E. coli TG1 by incubating eluted phage with bacteria for 30 minutes at 37°C. The E. coli are then plated on TYE plates containing 1% glucose and 100 µg/ml ampicillin. The resulting bacterial library is then rescued with delta gene 3 helper phage as described above to prepare phage for a subsequent round of selection. This process is then repeated for a total of 4 rounds of affinity purification with tube-washing increased to 20 times with PBS, 0.1% Tween-20 and 20 times with PBS for rounds 3 and 4.

Characterization of Binders. Eluted phage from the 3rd and 4th rounds of selection are used to infect E. coli HB 2151 and soluble scFv is produced (Marks, et al., 1991) from single colonies for assay. ELISAs are performed with microtitre plates

coated with either 10 pg/ml of the polypeptide of the present invention in 50 mM bicarbonate pH 9.6. Clones positive in ELISA are further characterized by PCR fingerprinting (see, e.g., PCT publication WO 92/01047) and then by sequencing. These ELISA positive clones may also be further characterized by techniques known  
5 in the art, such as, for example, epitope mapping, binding affinity, receptor signal transduction, ability to block or competitively inhibit antibody/antigen binding, and competitive agonistic or antagonistic activity.

Moreover, in another preferred method, the antibodies directed against the polypeptides of the present invention may be produced in plants. Specific methods are  
10 disclosed in US Patent Nos. 5,959,177, and 6,080,560, which are hereby incorporated in their entirety herein. The methods not only describe methods of expressing antibodies, but also the means of assembling foreign multimeric proteins in plants (i.e., antibodies, etc.), and the subsequent secretion of such antibodies from the plant.

### 15      EXAMPLE 33 – ASSAYS DETECTING STIMULATION OR INHIBITION OF    B CELL PROLIFERATION AND DIFFERENTIATION

Generation of functional humoral immune responses requires both soluble and cognate signaling between B-lineage cells and their microenvironment. Signals may impart a positive stimulus that allows a B-lineage cell to continue its programmed  
20 development, or a negative stimulus that instructs the cell to arrest its current developmental pathway. To date, numerous stimulatory and inhibitory signals have been found to influence B cell responsiveness including IL-2, IL-4, IL-5, IL-6, IL-7, IL10, IL-13, IL-14 and IL-15. Interestingly, these signals are by themselves weak effectors but can, in combination with various co-stimulatory proteins, induce  
25 activation, proliferation, differentiation, homing, tolerance and death among B cell populations.

One of the best studied classes of B-cell co-stimulatory proteins is the TNF-superfamily. Within this family CD40, CD27, and CD30 along with their respective ligands CD154, CD70, and CD153 have been found to regulate a variety of immune  
30 responses. Assays which allow for the detection and/or observation of the proliferation and differentiation of these B-cell populations and their precursors are valuable tools in determining the effects various proteins may have on these B-cell

populations in terms of proliferation and differentiation. Listed below are two assays designed to allow for the detection of the differentiation, proliferation, or inhibition of B-cell populations and their precursors.

In Vitro Assay- Purified polypeptides of the invention, or truncated forms thereof, is assessed for its ability to induce activation, proliferation, differentiation or inhibition and/or death in B-cell populations and their precursors. The activity of the polypeptides of the invention on purified human tonsillar B cells, measured qualitatively over the dose range from 0.1 to 10,000 ng/mL, is assessed in a standard B-lymphocyte co-stimulation assay in which purified tonsillar B cells are cultured in the presence of either formalin-fixed *Staphylococcus aureus* Cowan I (SAC) or immobilized anti-human IgM antibody as the priming agent. Second signals such as IL-2 and IL-15 synergize with SAC and IgM crosslinking to elicit B cell proliferation as measured by tritiated-thymidine incorporation. Novel synergizing agents can be readily identified using this assay. The assay involves isolating human tonsillar B cells by magnetic bead (MACS) depletion of CD3-positive cells. The resulting cell population is greater than 95% B cells as assessed by expression of CD45R(B220).

Various dilutions of each sample are placed into individual wells of a 96-well plate to which are added 10<sup>5</sup> B-cells suspended in culture medium (RPMI 1640 containing 10% FBS, 5 X 10<sup>-5</sup>M 2ME, 100U/ml penicillin, 10ug/ml streptomycin, and 10<sup>-5</sup> dilution of SAC) in a total volume of 150ul. Proliferation or inhibition is quantitated by a 20h pulse (1uCi/well) with 3H-thymidine (6.7 Ci/mM) beginning 72h post factor addition. The positive and negative controls are IL2 and medium respectively.

In Vivo Assay- BALB/c mice are injected (i.p.) twice per day with buffer only, or 2 mg/Kg of a polypeptide of the invention, or truncated forms thereof. Mice receive this treatment for 4 consecutive days, at which time they are sacrificed and various tissues and serum collected for analyses. Comparison of H&E sections from normal spleens and spleens treated with polypeptides of the invention identify the results of the activity of the polypeptides on spleen cells, such as the diffusion of periaxillary lymphatic sheaths, and/or significant increases in the nucleated cellularity of the red pulp regions, which may indicate the activation of the differentiation and proliferation of B-cell populations. Immunohistochemical studies using a B cell

marker, anti-CD45R(B220), are used to determine whether any physiological changes to splenic cells, such as splenic disorganization, are due to increased B-cell representation within loosely defined B-cell zones that infiltrate established T-cell regions.

5           Flow cytometric analyses of the spleens from mice treated with polypeptide is used to indicate whether the polypeptide specifically increases the proportion of ThB+, CD45R(B220)dull B cells over that which is observed in control mice.

          Likewise, a predicted consequence of increased mature B-cell representation in vivo is a relative increase in serum Ig titers. Accordingly, serum IgM and IgA  
10       levels are compared between buffer and polypeptide-treated mice.

          One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

#### 15                           EXAMPLE 34 – T CELL PROLIFERATION ASSAY

          A CD3-induced proliferation assay is performed on PBMCs and is measured by the uptake of 3H-thymidine. The assay is performed as follows. Ninety-six well plates are coated with 100 (l/well of mAb to CD3 (HIT3a, Pharmingen) or isotype-matched control mAb (B33.1) overnight at 4 degrees C (1 (g/ml in .05M bicarbonate  
20       buffer, pH 9.5), then washed three times with PBS. PBMC are isolated by F/H gradient centrifugation from human peripheral blood and added to quadruplicate wells (5 x 10<sup>4</sup>/well) of mAb coated plates in RPMI containing 10% FCS and P/S in the presence of varying concentrations of polypeptides of the invention (total volume 200 ul). Relevant protein buffer and medium alone are controls. After 48 hr. culture at 37  
25       degrees C, plates are spun for 2 min. at 1000 rpm and 100 (l of supernatant is removed and stored –20 degrees C for measurement of IL-2 (or other cytokines) if effect on proliferation is observed. Wells are supplemented with 100 ul of medium containing 0.5 uCi of 3H-thymidine and cultured at 37 degrees C for 18-24 hr. Wells are harvested and incorporation of 3H-thymidine used as a measure of proliferation.  
30       Anti-CD3 alone is the positive control for proliferation. IL-2 (100 U/ml) is also used as a control which enhances proliferation. Control antibody which does not induce

proliferation of T cells is used as the negative controls for the effects of polypeptides of the invention.

One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or  
5 antagonists of polynucleotides or polypeptides of the invention.

EXAMPLE 35 – EFFECT OF POLYPEPTIDES OF THE INVENTION ON THE  
EXPRESSION OF MHC CLASS II, COSTIMULATORY AND ADHESION  
MOLECULES AND CELL DIFFERENTIATION OF MONOCYTES AND  
10 MONOCYTE-DERIVED HUMAN DENDRITIC CELLS

Dendritic cells are generated by the expansion of proliferating precursors found in the peripheral blood: adherent PBMC or elutriated monocytic fractions are cultured for 7-10 days with GM-CSF (50 ng/ml) and IL-4 (20 ng/ml). These dendritic cells have the characteristic phenotype of immature cells (expression of CD1, CD80,  
15 CD86, CD40 and MHC class II antigens). Treatment with activating factors, such as TNF-, causes a rapid change in surface phenotype (increased expression of MHC class I and II, costimulatory and adhesion molecules, downregulation of FC(RII, upregulation of CD83). These changes correlate with increased antigen-presenting capacity and with functional maturation of the dendritic cells.

20 FACS analysis of surface antigens is performed as follows. Cells are treated 1-3 days with increasing concentrations of polypeptides of the invention or LPS (positive control), washed with PBS containing 1% BSA and 0.02 mM sodium azide, and then incubated with 1:20 dilution of appropriate FITC- or PE-labeled monoclonal antibodies for 30 minutes at 4 degrees C. After an additional wash, the labeled cells  
25 are analyzed by flow cytometry on a FACScan (Becton Dickinson).

Effect on the production of cytokines. Cytokines generated by dendritic cells, in particular IL-12, are important in the initiation of T-cell dependent immune responses. IL-12 strongly influences the development of Th1 helper T-cell immune response, and induces cytotoxic T and NK cell function. An ELISA is used to  
30 measure the IL-12 release as follows. Dendritic cells (106/ml) are treated with increasing concentrations of polypeptides of the invention for 24 hours. LPS (100 ng/ml) is added to the cell culture as positive control. Supernatants from the cell

cultures are then collected and analyzed for IL-12 content using commercial ELISA kit(e.g., R & D Systems (Minneapolis, MN)). The standard protocols provided with the kits are used.

Effect on the expression of MHC Class II, costimulatory and adhesion molecules. Three major families of cell surface antigens can be identified on monocytes: adhesion molecules, molecules involved in antigen presentation, and Fc receptor. Modulation of the expression of MHC class II antigens and other costimulatory molecules, such as B7 and ICAM-1, may result in changes in the antigen presenting capacity of monocytes and ability to induce T cell activation. Increase expression of Fc receptors may correlate with improved monocyte cytotoxic activity, cytokine release and phagocytosis.

FACS analysis is used to examine the surface antigens as follows. Monocytes are treated 1-5 days with increasing concentrations of polypeptides of the invention or LPS (positive control), washed with PBS containing 1% BSA and 0.02 mM sodium azide, and then incubated with 1:20 dilution of appropriate FITC- or PE-labeled monoclonal antibodies for 30 minutes at 4 degrees C. After an additional wash, the labeled cells are analyzed by flow cytometry on a FACScan (Becton Dickinson).

Monocyte activation and/or increased survival. Assays for molecules that activate (or alternatively, inactivate) monocytes and/or increase monocyte survival (or alternatively, decrease monocyte survival) are known in the art and may routinely be applied to determine whether a molecule of the invention functions as an inhibitor or activator of monocytes. Polypeptides, agonists, or antagonists of the invention can be screened using the three assays described below. For each of these assays, Peripheral blood mononuclear cells (PBMC) are purified from single donor leukopacks (American Red Cross, Baltimore, MD) by centrifugation through a Histopaque gradient (Sigma). Monocytes are isolated from PBMC by counterflow centrifugal elutriation.

Monocyte Survival Assay. Human peripheral blood monocytes progressively lose viability when cultured in absence of serum or other stimuli. Their death results from internally regulated process (apoptosis). Addition to the culture of activating factors, such as TNF-alpha dramatically improves cell survival and prevents DNA fragmentation. Propidium iodide (PI) staining is used to measure apoptosis as follows.

Monocytes are cultured for 48 hours in polypropylene tubes in serum-free medium (positive control), in the presence of 100 ng/ml TNF-alpha (negative control), and in the presence of varying concentrations of the compound to be tested. Cells are suspended at a concentration of  $2 \times 10^6$ /ml in PBS containing PI at a final  
5 concentration of 5 (g/ml, and then incubated at room temperature for 5 minutes before FACScan analysis. PI uptake has been demonstrated to correlate with DNA fragmentation in this experimental paradigm.

Effect on cytokine release. An important function of monocytes/macrophages is their regulatory activity on other cellular populations of the immune system through  
10 the release of cytokines after stimulation. An ELISA to measure cytokine release is performed as follows. Human monocytes are incubated at a density of  $5 \times 10^5$  cells/ml with increasing concentrations of the a polypeptide of the invention and under the same conditions, but in the absence of the polypeptide. For IL-12 production, the cells are primed overnight with IFN (100 U/ml) in presence of a polypeptide of the  
15 invention. LPS (10 ng/ml) is then added. Conditioned media are collected after 24h and kept frozen until use. Measurement of TNF-alpha, IL-10, MCP-1 and IL-8 is then performed using a commercially available ELISA kit(e.g., R & D Systems (Minneapolis, MN)) and applying the standard protocols provided with the kit.

Oxidative burst. Purified monocytes are plated in 96-w plate at  $2 \times 10^5$   
20 cell/well. Increasing concentrations of polypeptides of the invention are added to the wells in a total volume of 0.2 ml culture medium (RPMI 1640 + 10% FCS, glutamine and antibiotics). After 3 days incubation, the plates are centrifuged and the medium is removed from the wells. To the macrophage monolayers, 0.2 ml per well of phenol red solution (140 mM NaCl, 10 mM potassium phosphate buffer pH 7.0, 5.5 mM  
25 dextrose, 0.56 mM phenol red and 19 U/ml of HRPO) is added, together with the stimulant (200 nM PMA). The plates are incubated at 37(C for 2 hours and the reaction is stopped by adding 20  $\mu$ l 1N NaOH per well. The absorbance is read at 610 nm. To calculate the amount of H<sub>2</sub>O<sub>2</sub> produced by the macrophages, a standard curve of a H<sub>2</sub>O<sub>2</sub> solution of known molarity is performed for each experiment.

30 One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

EXAMPLE 36 – BIOLOGICAL EFFECTS OF HUMAN PHOSPHATASE  
POLYPEPTIDES OF THE INVENTION

Astrocyte and Neuronal Assays

5           Recombinant polypeptides of the invention, expressed in *Escherichia coli* and purified as described above, can be tested for activity in promoting the survival, neurite outgrowth, or phenotypic differentiation of cortical neuronal cells and for inducing the proliferation of glial fibrillary acidic protein immunopositive cells, astrocytes. The selection of cortical cells for the bioassay is based on the prevalent  
10       expression of FGF-1 and FGF-2 in cortical structures and on the previously reported enhancement of cortical neuronal survival resulting from FGF-2 treatment. A thymidine incorporation assay, for example, can be used to elucidate a polypeptide of the invention's activity on these cells.

          Moreover, previous reports describing the biological effects of FGF-2 (basic  
15       FGF) on cortical or hippocampal neurons in vitro have demonstrated increases in both neuron survival and neurite outgrowth (Walicke et al., "Fibroblast growth factor promotes survival of dissociated hippocampal neurons and enhances neurite extension." *Proc. Natl. Acad. Sci. USA* 83:3012-3016. (1986), assay herein incorporated by reference in its entirety). However, reports from experiments done on  
20       PC-12 cells suggest that these two responses are not necessarily synonymous and may depend on not only which FGF is being tested but also on which receptor(s) are expressed on the target cells. Using the primary cortical neuronal culture paradigm, the ability of a polypeptide of the invention to induce neurite outgrowth can be compared to the response achieved with FGF-2 using, for example, a thymidine  
25       incorporation assay.

Fibroblast and endothelial cell assays

          Human lung fibroblasts are obtained from Clonetics (San Diego, CA) and maintained in growth media from Clonetics. Dermal microvascular endothelial cells are obtained from Cell Applications (San Diego, CA). For proliferation assays, the  
30       human lung fibroblasts and dermal microvascular endothelial cells can be cultured at 5,000 cells/well in a 96-well plate for one day in growth medium. The cells are then incubated for one day in 0.1% BSA basal medium. After replacing the medium with

fresh 0.1% BSA medium, the cells are incubated with the test proteins for 3 days. Alamar Blue (Alamar Biosciences, Sacramento, CA) is added to each well to a final concentration of 10%. The cells are incubated for 4 hr. Cell viability is measured by reading in a CytoFluor fluorescence reader. For the PGE2 assays, the human lung  
 5 fibroblasts are cultured at 5,000 cells/well in a 96-well plate for one day. After a medium change to 0.1% BSA-basal medium, the cells are incubated with FGF-2 or polypeptides of the invention with or without IL-1( for 24 hours. The supernatants are collected and assayed for PGE2 by EIA kit (Cayman, Ann Arbor, MI). For the IL-6 assays, the human lung fibroblasts are cultured at 5,000 cells/well in a 96-well plate  
 10 for one day. After a medium change to 0.1% BSA basal medium, the cells are incubated with FGF-2 or with or without polypeptides of the invention IL-1( for 24 hours. The supernatants are collected and assayed for IL-6 by ELISA kit (Endogen, Cambridge, MA).

Human lung fibroblasts are cultured with FGF-2 or polypeptides of the  
 15 invention for 3 days in basal medium before the addition of Alamar Blue to assess effects on growth of the fibroblasts. FGF-2 should show a stimulation at 10 - 2500 ng/ml which can be used to compare stimulation with polypeptides of the invention.

#### Parkinson Models

The loss of motor function in Parkinson's disease is attributed to a deficiency  
 20 of striatal dopamine resulting from the degeneration of the nigrostriatal dopaminergic projection neurons. An animal model for Parkinson's that has been extensively characterized involves the systemic administration of 1-methyl-4 phenyl 1,2,3,6-tetrahydropyridine (MPTP). In the CNS, MPTP is taken-up by astrocytes and catabolized by monoamine oxidase B to 1-methyl-4-phenyl pyridine (MPP+) and  
 25 released. Subsequently, MPP+ is actively accumulated in dopaminergic neurons by the high-affinity reuptake transporter for dopamine. MPP+ is then concentrated in mitochondria by the electrochemical gradient and selectively inhibits nicotinamide adenine disphosphate: ubiquinone oxidoreductionase (complex I), thereby interfering with electron transport and eventually generating oxygen radicals.

30 It has been demonstrated in tissue culture paradigms that FGF-2 (basic FGF) has trophic activity towards nigral dopaminergic neurons (Ferrari et al., Dev. Biol. 1989). Recently, Dr. Unsicker's group has demonstrated that administering FGF-2 in

gel foam implants in the striatum results in the near complete protection of nigral dopaminergic neurons from the toxicity associated with MPTP exposure (Otto and Unsicker, J. Neuroscience, 1990).

Based on the data with FGF-2, polypeptides of the invention can be evaluated to determine whether it has an action similar to that of FGF-2 in enhancing dopaminergic neuronal survival in vitro and it can also be tested in vivo for protection of dopaminergic neurons in the striatum from the damage associated with MPTP treatment. The potential effect of a polypeptide of the invention is first examined in vitro in a dopaminergic neuronal cell culture paradigm. The cultures are prepared by dissecting the midbrain floor plate from gestation day 14 Wistar rat embryos. The tissue is dissociated with trypsin and seeded at a density of 200,000 cells/cm<sup>2</sup> on polyorthinine-laminin coated glass coverslips. The cells are maintained in Dulbecco's Modified Eagle's medium and F12 medium containing hormonal supplements (N1). The cultures are fixed with paraformaldehyde after 8 days in vitro and are processed for tyrosine hydroxylase, a specific marker for dopaminergic neurons, immunohistochemical staining. Dissociated cell cultures are prepared from embryonic rats. The culture medium is changed every third day and the factors are also added at that time.

Since the dopaminergic neurons are isolated from animals at gestation day 14, a developmental time which is past the stage when the dopaminergic precursor cells are proliferating, an increase in the number of tyrosine hydroxylase immunopositive neurons would represent an increase in the number of dopaminergic neurons surviving in vitro. Therefore, if a polypeptide of the invention acts to prolong the survival of dopaminergic neurons, it would suggest that the polypeptide may be involved in Parkinson's Disease.

One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

EXAMPLE 37 – THE EFFECT OF THE HUMAN PHOSPHATASE  
POLYPEPTIDES OF THE INVENTION ON THE GROWTH OF  
VASCULAR ENDOTHELIAL CELLS

On day 1, human umbilical vein endothelial cells (HUVEC) are seeded at 2-  
5 5x10<sup>4</sup> cells/35 mm dish density in M199 medium containing 4% fetal bovine serum  
(FBS), 16 units/ml heparin, and 50 units/ml endothelial cell growth supplements  
(ECGS, Biotechnology, Inc.). On day 2, the medium is replaced with M199 containing  
10% FBS, 8 units/ml heparin. A polypeptide having the amino acid sequence of SEQ  
ID NO:2, and positive controls, such as VEGF and basic FGF (bFGF) are added, at  
10 varying concentrations. On days 4 and 6, the medium is replaced. On day 8, cell  
number is determined with a Coulter Counter.

An increase in the number of HUVEC cells indicates that the polypeptide of  
the invention may proliferate vascular endothelial cells.

One skilled in the art could easily modify the exemplified studies to test the  
15 activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or  
antagonists of polynucleotides or polypeptides of the invention.

EXAMPLE 39 – STIMULATION OF ENDOTHELIAL MIGRATION

This example will be used to explore the possibility that a polypeptide of the  
20 invention may stimulate lymphatic endothelial cell migration.

Endothelial cell migration assays are performed using a 48 well  
microchemotaxis chamber (Neuroprobe Inc., Cabin John, MD; Falk, W., et al., J.  
Immunological Methods 1980;33:239-247). Polyvinylpyrrolidone-free polycarbonate  
filters with a pore size of 8  $\mu$ m (Nucleopore Corp. Cambridge, MA) are coated with  
25 0.1% gelatin for at least 6 hours at room temperature and dried under sterile air. Test  
substances are diluted to appropriate concentrations in M199 supplemented with  
0.25% bovine serum albumin (BSA), and 25  $\mu$ l of the final dilution is placed in the  
lower chamber of the modified Boyden apparatus. Subconfluent, early passage (2-6)  
HUVEC or BMEC cultures are washed and trypsinized for the minimum time  
30 required to achieve cell detachment. After placing the filter between lower and upper  
chamber, 2.5 x 10<sup>5</sup> cells suspended in 50  $\mu$ l M199 containing 1% FBS are seeded in  
the upper compartment. The apparatus is then incubated for 5 hours at 37°C in a

humidified chamber with 5% CO<sub>2</sub> to allow cell migration. After the incubation period, the filter is removed and the upper side of the filter with the non-migrated cells is scraped with a rubber policeman. The filters are fixed with methanol and stained with a Giemsa solution (Diff-Quick, Baxter, McGraw Park, IL). Migration is  
5 quantified by counting cells of three random high-power fields (40x) in each well, and all groups are performed in quadruplicate.

One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

10

#### EXAMPLE 40 – EFFECT OF HUMAN PHOSPHATASE POLYPEPTIDES OF THE INVENTION ON CORD FORMATION IN ANGIOGENESIS

Another step in angiogenesis is cord formation, marked by differentiation of endothelial cells. This bioassay measures the ability of microvascular endothelial cells  
15 to form capillary-like structures (hollow structures) when cultured in vitro.

CADMEC (microvascular endothelial cells) are purchased from Cell Applications, Inc. as proliferating (passage 2) cells and are cultured in Cell Applications' CADMEC Growth Medium and used at passage 5. For the in vitro angiogenesis assay, the wells of a 48-well cell culture plate are coated with Cell  
20 Applications' Attachment Factor Medium (200 µl/well) for 30 min. at 37°C. CADMEC are seeded onto the coated wells at 7,500 cells/well and cultured overnight in Growth Medium. The Growth Medium is then replaced with 300 µg Cell Applications' Chord Formation Medium containing control buffer or a polypeptide of the invention (0.1 to 100 ng/ml) and the cells are cultured for an additional 48 hr. The  
25 numbers and lengths of the capillary-like chords are quantitated through use of the Boeckeler VIA-170 video image analyzer. All assays are done in triplicate.

Commercial (R&D) VEGF (50 ng/ml) is used as a positive control. b-esteradiol (1 ng/ml) is used as a negative control. The appropriate buffer (without protein) is also utilized as a control.

30 One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

## EXAMPLE 41 – RESCUE OF ISCHEMIA IN RABBIT LOWER LIMB MODEL

To study the in vivo effects of polynucleotides and polypeptides of the invention on ischemia, a rabbit hindlimb ischemia model is created by surgical removal of one femoral arteries as described previously (Takeshita et al., Am J. Pathol 147:1649-1660 (1995)). The excision of the femoral artery results in retrograde propagation of thrombus and occlusion of the external iliac artery. Consequently, blood flow to the ischemic limb is dependent upon collateral vessels originating from the internal iliac artery (Takeshita et al. Am J. Pathol 147:1649-1660 (1995)). An interval of 10 days is allowed for post-operative recovery of rabbits and development of endogenous collateral vessels. At 10 day post-operatively (day 0), after performing a baseline angiogram, the internal iliac artery of the ischemic limb is transfected with 500 mg naked expression plasmid containing a polynucleotide of the invention by arterial gene transfer technology using a hydrogel-coated balloon catheter as described (Riessen et al. Hum Gene Ther. 4:749-758 (1993); Leclerc et al. J. Clin. Invest. 90: 936-944 (1992)). When a polypeptide of the invention is used in the treatment, a single bolus of 500 mg polypeptide of the invention or control is delivered into the internal iliac artery of the ischemic limb over a period of 1 min. through an infusion catheter. On day 30, various parameters are measured in these rabbits: (a) BP ratio - The blood pressure ratio of systolic pressure of the ischemic limb to that of normal limb; (b) Blood Flow and Flow Reserve - Resting FL: the blood flow during undilated condition and Max FL: the blood flow during fully dilated condition (also an indirect measure of the blood vessel amount) and Flow Reserve is reflected by the ratio of max FL: resting FL; (c) Angiographic Score - This is measured by the angiogram of collateral vessels. A score is determined by the percentage of circles in an overlaying grid that with crossing opacified arteries divided by the total number in the rabbit thigh; (d) Capillary density - The number of collateral capillaries determined in light microscopic sections taken from hindlimbs.

One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

#### EXAMPLE 42 – RAT CORNEAL WOUND HEALING MODEL

This animal model shows the effect of a polypeptide of the invention on neovascularization. The experimental protocol includes:

- 5 a) Making a 1-1.5 mm long incision from the center of cornea into the stromal layer.
- b) Inserting a spatula below the lip of the incision facing the outer corner of the eye.
- c) Making a pocket (its base is 1-1.5 mm from the edge of the eye).
- 10 d) Positioning a pellet, containing 50ng- 5ug of a polypeptide of the invention, within the pocket.
- e) Treatment with a polypeptide of the invention can also be applied topically to the corneal wounds in a dosage range of 20mg - 500mg (daily treatment for five days).

15 One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

#### EXAMPLE 43 – SUPPRESSION OF TNF ALPHA-INDUCED ADHESION MOLECULE EXPRESSION BY A POLYPEPTIDE OF THE INVENTION

20 The recruitment of lymphocytes to areas of inflammation and angiogenesis involves specific receptor-ligand interactions between cell surface adhesion molecules (CAMs) on lymphocytes and the vascular endothelium. The adhesion process, in both normal and pathological settings, follows a multi-step cascade that involves intercellular adhesion molecule-1 (ICAM-1), vascular cell adhesion molecule-1  
25 (VCAM-1), and endothelial leukocyte adhesion molecule-1 (E-selectin) expression on endothelial cells (EC). The expression of these molecules and others on the vascular endothelium determines the efficiency with which leukocytes may adhere to the local vasculature and extravasate into the local tissue during the development of an inflammatory response. The local concentration of cytokines and growth factor  
30 participate in the modulation of the expression of these CAMs.

Tumor necrosis factor alpha (TNF- $\alpha$ ), a potent proinflammatory cytokine, is a stimulator of all three CAMs on endothelial cells and may be involved in a wide variety of inflammatory responses, often resulting in a pathological outcome.

The potential of a polypeptide of the invention to mediate a suppression of  
 5 TNF- $\alpha$  induced CAM expression can be examined. A modified ELISA assay which uses ECs as a solid phase absorbent is employed to measure the amount of CAM expression on TNF- $\alpha$  treated ECs when co-stimulated with a member of the FGF family of proteins.

To perform the experiment, human umbilical vein endothelial cell (HUVEC)  
 10 cultures are obtained from pooled cord harvests and maintained in growth medium (EGM-2; Clonetics, San Diego, CA) supplemented with 10% FCS and 1% penicillin/streptomycin in a 37 degree C humidified incubator containing 5% CO<sub>2</sub>. HUVECs are seeded in 96-well plates at concentrations of  $1 \times 10^4$  cells/well in EGM medium at 37 degree C for 18-24 hrs or until confluent. The monolayers are  
 15 subsequently washed 3 times with a serum-free solution of RPMI-1640 supplemented with 100 U/ml penicillin and 100 mg/ml streptomycin, and treated with a given cytokine and/or growth factor(s) for 24 h at 37 degree C. Following incubation, the cells are then evaluated for CAM expression.

Human Umbilical Vein Endothelial cells (HUVECs) are grown in a standard  
 20 96 well plate to confluence. Growth medium is removed from the cells and replaced with 90  $\mu$ l of 199 Medium (10% FBS). Samples for testing and positive or negative controls are added to the plate in triplicate (in 10  $\mu$ l volumes). Plates are incubated at 37 degree C for either 5 h (selectin and integrin expression) or 24 h (integrin expression only). Plates are aspirated to remove medium and 100  $\mu$ l of 0.1%  
 25 paraformaldehyde-PBS(with Ca<sup>++</sup> and Mg<sup>++</sup>) is added to each well. Plates are held at 4oC for 30 min.

Fixative is then removed from the wells and wells are washed 1X with PBS(+Ca,Mg)+0.5% BSA and drained. Do not allow the wells to dry. Add 10  $\mu$ l of diluted primary antibody to the test and control wells. Anti-ICAM-1-Biotin, Anti-  
 30 VCAM-1-Biotin and Anti-E-selectin-Biotin are used at a concentration of 10  $\mu$ g/ml (1:10 dilution of 0.1 mg/ml stock antibody). Cells are incubated at 37oC for 30 min. in a humidified environment. Wells are washed X3 with PBS(+Ca,Mg)+0.5% BSA.

Then add 20 µl of diluted ExtrAvidin-Alkaline Phosphatase (1:5,000 dilution) to each well and incubated at 37°C for 30 min. Wells are washed X3 with PBS(+Ca,Mg)+0.5% BSA. 1 tablet of p-Nitrophenol Phosphate pNPP is dissolved in 5 ml of glycine buffer (pH 10.4). 100 µl of pNPP substrate in glycine buffer is added to each test well. Standard wells in triplicate are prepared from the working dilution of the ExtrAvidin-Alkaline Phosphatase in glycine buffer: 1:5,000.(100) > 10<sup>-0.5</sup> > 10<sup>-1</sup> > 10<sup>-1.5</sup>. 5 µl of each dilution is added to triplicate wells and the resulting AP content in each well is 5.50 ng, 1.74 ng, 0.55 ng, 0.18 ng. 100 µl of pNPP reagent must then be added to each of the standard wells. The plate must be incubated at 37°C for 4h. A volume of 50 µl of 3M NaOH is added to all wells. The results are quantified on a plate reader at 405 nm. The background subtraction option is used on blank wells filled with glycine buffer only. The template is set up to indicate the concentration of AP-conjugate in each standard well [ 5.50 ng; 1.74 ng; 0.55 ng; 0.18 ng]. Results are indicated as amount of bound AP-conjugate in each sample.

15

EXAMPLE 44 – METHOD OF CREATING N- AND C-TERMINAL DELETION  
MUTANTS CORRESPONDING TO THE HUMAN PHOSPHATASE  
POLYPEPTIDES OF THE PRESENT INVENTION

As described elsewhere herein, the present invention encompasses the creation of N- and C-terminal deletion mutants, in addition to any combination of N- and C-terminal deletions thereof, corresponding to the human phosphatase polypeptides of the present invention. A number of methods are available to one skilled in the art for creating such mutants. Such methods may include a combination of PCR amplification and gene cloning methodology. Although one of skill in the art of molecular biology, through the use of the teachings provided or referenced herein, and/or otherwise known in the art as standard methods, could readily create each deletion mutant of the present invention, exemplary methods are described below using specific BMY\_HPP13 deletions as examples.

Briefly, using the isolated cDNA clone encoding the full-length human BMY\_HPP13 phosphatase polypeptide sequence (as described elsewhere herein, for example), appropriate primers of about 15-25 nucleotides derived from the desired 5' and 3' positions of SEQ ID NO:1 may be designed to PCR amplify, and subsequently

clone, the intended N- and/or C-terminal deletion mutant. Such primers could comprise, for example, an initiation and stop codon for the 5' and 3' primer, respectively. Such primers may also comprise restriction sites to facilitate cloning of the deletion mutant post amplification. Moreover, the primers may comprise additional sequences, such as, for example, flag-tag sequences, kozac sequences, or other sequences discussed and/or referenced herein.

For example, in the case of the E18 to E246 BMY\_HPP13 N-terminal deletion mutant, the following primers could be used to amplify a cDNA fragment corresponding to this deletion mutant:

5' Primer	5'-GCAGCA <u>GCGGCCGC</u> GAGGCAAAACCCCGGGCCACATGG -3' (SEQ ID NO:35) <i>NotI</i>
3' Primer	5'- GCAGCA <u>GTCGAC</u> TTCTGATGTAAAGCATTGACTAC -3' (SEQ ID NO:36) <i>Sall</i>

For example, in the case of the M1 to G213 BMY\_HPP13 C-terminal deletion mutant, the following primers could be used to amplify a cDNA fragment corresponding to this deletion mutant:

5' Primer	5'- GCAGCA <u>GCGGCCGC</u> ATG GTTGTAGATTTCTGGACTTGGG -3' (SEQ ID NO:37) <i>NotI</i>
3' Primer	5'- GCAGCA <u>GTCGAC</u> CCCACTCCCCTTAGCCTCTTTTGCC -3' (SEQ ID NO:38) <i>Sall</i>

Representative PCR amplification conditions are provided below, although the skilled artisan would appreciate that other conditions may be required for efficient amplification. A 100 ul PCR reaction mixture may be prepared using 10ng of the template DNA (cDNA clone of Human phosphatase polypeptides), 200 uM 4dNTPs, 1uM primers, 0.25U Taq DNA polymerase (PE), and standard Taq DNA polymerase buffer. Typical PCR cycling condition are as follows:

20-25 cycles: 45 sec, 93 degrees

2 min, 50 degrees

2 min, 72 degrees

1 cycle: 10 min, 72 degrees

After the final extension step of PCR, 5U Klenow Fragment may be added and incubated for 15 min at 30 degrees.

Upon digestion of the fragment with the NotI and SalI restriction enzymes, the  
5 fragment could be cloned into an appropriate expression and/or cloning vector which  
has been similarly digested (e.g., pSport1, among others). . The skilled artisan would  
appreciate that other plasmids could be equally substituted, and may be desirable in  
certain circumstances. The digested fragment and vector are then ligated using a DNA  
ligase, and then used to transform competent E.coli cells using methods provided  
10 herein and/or otherwise known in the art.

The 5' primer sequence for amplifying any additional N-terminal deletion  
mutants may be determined by reference to the following formula:  $(S+(X * 3))$  to  
 $((S+(X * 3))+25)$ , wherein 'S' is equal to the nucleotide position of the initiating start  
codon of the human BMY\_HPP13 phosphatase gene (SEQ ID NO:1), and 'X' is equal  
15 to the most N-terminal amino acid of the intended N-terminal deletion mutant. The  
first term will provide the start 5' nucleotide position of the 5' primer, while the  
second term will provide the end 3' nucleotide position of the 5' primer  
corresponding to sense strand SEQ ID NO:1. Once the corresponding nucleotide  
positions of the primer are determined, the final nucleotide sequence may be created  
20 by the addition of applicable restriction site sequences to the 5' end of the sequence,  
for example. As referenced herein, the addition of other sequences to the 5' primer  
may be desired in certain circumstances (e.g., kozac sequences, etc.).

The 3' primer sequence for amplifying any additional N-terminal deletion  
mutants may be determined by reference to the following formula:  $(S+(X * 3))$  to  
25  $((S+(X * 3))-25)$ , wherein 'S' is equal to the nucleotide position of the initiating start  
codon of the human BMY\_HPP13 phosphatase genes (SEQ SEQ ID NO:1), and 'X'  
is equal to the most C-terminal amino acid of the intended N-terminal deletion  
mutant. The first term will provide the start 5' nucleotide position of the 3' primer,  
while the second term will provide the end 3' nucleotide position of the 3' primer  
30 corresponding to the anti-sense strand of SEQ SEQ ID NO:1, respectively. Once the  
corresponding nucleotide positions of the primer are determined, the final nucleotide  
sequence may be created by the addition of applicable restriction site sequences to the

5' end of the sequence, for example. As referenced herein, the addition of other sequences to the 3' primer may be desired in certain circumstances (e.g., stop codon sequences, etc.). The skilled artisan would appreciate that modifications of the above nucleotide positions may be necessary for optimizing PCR amplification.

5       The same general formulas provided above may be used in identifying the 5' and 3' primer sequences for amplifying any C-terminal deletion mutant of the present invention. Moreover, the same general formulas provided above may be used in identifying the 5' and 3' primer sequences for amplifying any combination of N-terminal and C-terminal deletion mutant of the present invention. The skilled artisan  
10       would appreciate that modifications of the above nucleotide positions may be necessary for optimizing PCR amplification.

One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

15

#### EXAMPLE 45 – SITE DIRECTED/SITE-SPECIFIC MUTAGENESIS

*In vitro* site-directed mutagenesis is an invaluable technique for studying protein structure-function relationships and gene expression, for example, as well as for vector modification. Site-directed mutagenesis can also be used for creating any of  
20       one or more of the mutants of the present invention, particularly the conservative and/or non-conservative amino acid substitution mutants of the present invention. Approaches utilizing single stranded DNA (ssDNA) as the template have been reported (e.g., T.A. Kunkel et al., 1985, *Proc. Natl. Acad. Sci. USA*), 82:488-492; M.A. Vandeyar et al., 1988, *Gene*, 65(1):129-133; M. Sugimoto et al., 1989, *Anal.*  
25       *Biochem.*, 179(2):309-311; and J.W. Taylor et al., 1985, *Nuc. Acids. Res.*, 13(24):8765-8785).

The use of PCR in site-directed mutagenesis accomplishes strand separation by using a denaturing step to separate the complementary strands and to allow efficient polymerization of the PCR primers. PCR site-directed mutagenesis methods  
30       thus permit site specific mutations to be incorporated in virtually any double stranded plasmid, thus eliminating the need for re-subcloning into M13-based bacteriophage vectors or single-stranded rescue. (M.P. Weiner et al., 1995, *Molecular Biology*:

*Current Innovations and Future Trends*, Eds. A.M. Griffin and H.G. Griffin, Horizon Scientific Press, Norfolk, UK; and C. Papworth et al., 1996, *Strategies*, 9(3):3-4).

A protocol for performing site-directed mutagenesis, particularly employing the QuikChange™ site-directed mutagenesis kit (Stratagene, La Jolla, CA; U.S. Patent Nos. 5,789,166 and 5,923,419) is provided for making point mutations, to switch or substitute amino acids, and to delete or insert single or multiple amino acids in the RATL1d6 amino acid sequence of this invention.

#### Primer Design

For primer design using this protocol, the mutagenic oligonucleotide primers are designed individually according to the desired mutation. The following considerations should be made for designing mutagenic primers: 1) Both of the mutagenic primers must contain the desired mutation and anneal to the same sequence on opposite strands of the plasmid; 2) Primers should be between 25 and 45 bases in length, and the melting temperature ( $T_m$ ) of the primers should be greater than, or equal to, 78°C. The following formula is commonly used for estimating the  $T_m$  of primers:  $T = 81.5 + 0.41 (\%GC) - 675/N - \%mismatch$ . For calculating  $T_m$ ,  $N$  is the primer length in bases; and values for %GC and % mismatch are whole numbers. For calculating  $T_m$  for primers intended to introduce insertions or deletions, a modified version of the above formula is employed:  $T = 81.5 + 0.41 (\%GC) - 675/N$ , where  $N$  does not include the bases which are being inserted or deleted; 3) The desired mutation (deletion or insertion) should be in the middle of the primer with approximately 10-15 bases of correct sequence on both sides; 4) The primers optimally should have a minimum GC content of 40%, and should terminate in one or more C or G bases; 5) Primers need not be 5'-phosphorylated, but must be purified either by fast polynucleotide liquid chromatography (FPLC) or by polyacrylamide gel electrophoresis (PAGE). Failure to purify the primers results in a significant decrease in mutation efficiency; and 6) It is important that primer concentration is in excess. It is suggested to vary the amount of template while keeping the concentration of the primers constantly in excess (QuikChange™ Site-Directed Mutagenesis Kit, Stratagene, La Jolla, CA).

### Protocol for Setting Up the Reactions

Using the above-described primer design, two complimentary oligonucleotides containing the desired mutation, flanked by unmodified nucleic acid sequence, are synthesized. The resulting oligonucleotide primers are purified.

5           A control reaction is prepared using 5 µl 10x reaction buffer (100mM KCl; 100mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; 200mM Tris-HCl, pH 8.8; 20mM MgSO<sub>4</sub>; 1% Triton® X-100; 1 mg/ml nuclease-free bovine serum albumin, BSA); 2 µl (10ng) of pWhitescript™, 4.5-kb control plasmid (5 ng/µl); 1.25 µl (125 ng) of oligonucleotide control primer #1 (34-mer, 100 ng/µl); 1.25 µl (125 ng) of oligonucleotide control primer #2 (34-mer, 100 ng/µl); 1 µl of dNTP mix; double distilled H<sub>2</sub>O; to a final volume of 50 µl. Thereafter, 1 µl of DNA polymerase (*PfuTurbo*® DNA Polymerase, Stratagene), (2.5U/µl) is added. *PfuTurbo*® DNA Polymerase is stated to have 6-fold higher fidelity in DNA synthesis than does *Taq* polymerase. To maximize temperature cycling performance, use of thin-walled test tubes is suggested to ensure optimum  
10           contact with the heating blocks of the temperature cycler.  
15

          The sample reaction is prepared by combining 5 µl of 10x reaction buffer; x µl (5-50 ng) of dsDNA template; x µl (125 ng) of oligonucleotide primer #1; x µl (5-50 ng) of dsDNA template; x µl (125 ng) of oligonucleotide primer #2; 1 µl of dNTP mix; and ddH<sub>2</sub>O to a final volume of 50 µl. Thereafter, 1 µl of DNA polymerase (*PfuTurbo* DNA Polymerase, Stratagene), (2.5U/µl) is added.  
20

          It is suggested that if the thermal cycler does not have a hot-top assembly, each reaction should be overlaid with approximately 30 µl of mineral oil.

### Cycling the Reactions

Each reaction is cycled using the following cycling parameters:

25

SegmentCyclesTemperatureTime:

1195°C	30 seconds
212-1895°C	30 seconds
55°C	1 minute
30       68°C	2 minutes/kb of plasmid length

For the control reaction, a 12-minute extension time is used and the reaction is run for 12 cycles. Segment 2 of the above cycling parameters is adjusted in accordance with the type of mutation desired. For example, for point mutations, 12 cycles are used; for single amino acid changes, 16 cycles are used; and for multiple amino acid deletions or insertions, 18 cycles are used. Following the temperature cycling, the reaction is placed on ice for 2 minutes to cool the reaction to  $\leq 37^{\circ}\text{C}$ .

#### Digesting the Products and Transforming Competent Cells

One  $\mu\text{l}$  of the *DpnI* restriction enzyme (10U/ $\mu\text{l}$ ) is added directly (below mineral oil overlay) to each amplification reaction using a small, pointed pipette tip. The reaction mixture is gently and thoroughly mixed by pipetting the solution up and down several times. The reaction mixture is then centrifuged for 1 minute in a microcentrifuge. Immediately thereafter, each reaction is incubated at  $37^{\circ}\text{C}$  for 1 hour to digest the parental (i.e., the non-mutated) supercoiled dsDNA.

Competent cells (i.e., XL1-Blue supercompetent cells, Stratagene) are thawed gently on ice. For each control and sample reaction to be transformed, 50  $\mu\text{l}$  of the supercompetent cells are aliquotted to a prechilled test tube (Falcon 2059 polypropylene). Next, 1  $\mu\text{l}$  of the *DpnI*-digested DNA is transferred from the control and the sample reactions to separate aliquots of the supercompetent cells. The transformation reactions are gently swirled to mix and incubated for 30 minutes on ice. Thereafter, the transformation reactions are heat-pulsed for 45 seconds at  $42^{\circ}\text{C}$  for 2 minutes.

0.5 ml of NZY+ broth, preheated to  $42^{\circ}\text{C}$  is added to the transformation reactions which are then incubated at  $37^{\circ}\text{C}$  for 1 hour with shaking at 225-250 rpm. An aliquot of each transformation reaction is plated on agar plates containing the appropriate antibiotic for the vector. For the mutagenesis and transformation controls, cells are spread on LB-ampicillin agar plates containing 80  $\mu\text{g/ml}$  of X-gal and 20mM IPTG. Transformation plates are incubated for  $>16$  hours at  $37^{\circ}\text{C}$ .

One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous

modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

5 The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

TABLE IV

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1	N	MET	1	72.945	18.091	11.490
2	CA	MET	1	73.488	17.114	10.537
3	C	MET	1	72.508	16.864	9.399
4	O	MET	1	72.833	17.113	8.231
5	CB	MET	1	73.807	15.801	11.239
6	CG	MET	1	74.833	16.007	12.348
7	SD	MET	1	75.352	14.508	13.214
8	CE	MET	1	76.062	13.606	11.817
9	N	VAL	2	71.259	16.615	9.764
10	CA	VAL	2	70.217	16.290	8.776
11	C	VAL	2	69.821	17.499	7.927
12	O	VAL	2	69.601	17.356	6.717
13	CB	VAL	2	68.999	15.776	9.538
14	CG1	VAL	2	67.822	15.491	8.611
15	CG2	VAL	2	69.355	14.536	10.349
16	N	VAL	3	70.028	18.684	8.480
17	CA	VAL	3	69.754	19.919	7.745
18	C	VAL	3	70.940	20.312	6.864
19	O	VAL	3	70.744	20.866	5.779
20	CB	VAL	3	69.524	21.030	8.766
21	CG1	VAL	3	68.967	22.281	8.095
22	CG2	VAL	3	68.597	20.580	9.888
23	N	ASP	4	72.108	19.788	7.199
24	CA	ASP	4	73.343	20.194	6.524
25	C	ASP	4	73.647	19.267	5.356
26	O	ASP	4	74.338	19.648	4.405
27	CB	ASP	4	74.488	20.127	7.533
28	CG	ASP	4	74.207	21.040	8.726
29	OD1	ASP	4	74.546	22.210	8.639
30	OD2	ASP	4	73.590	20.569	9.677
31	N	PHE	5	73.103	18.064	5.428
32	CA	PHE	5	73.206	17.124	4.314
33	C	PHE	5	71.920	17.085	3.492
34	O	PHE	5	71.848	16.348	2.500
35	CB	PHE	5	73.535	15.745	4.870
36	CG	PHE	5	74.910	15.666	5.528
37	CD1	PHE	5	76.045	16.000	4.800
38	CD2	PHE	5	75.030	15.255	6.850

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
39	CE1	PHE	5	77.297	15.933	5.394
40	CE2	PHE	5	76.283	15.189	7.445
41	CZ	PHE	5	77.416	15.528	6.717
42	N	TRP	6	70.926	17.847	3.935
43	CA	TRP	6	69.634	17.991	3.248
44	C	TRP	6	68.945	16.633	3.114
45	O	TRP	6	68.446	16.254	2.050
46	CB	TRP	6	69.912	18.621	1.886
47	CG	TRP	6	68.718	19.214	1.175
48	CD1	TRP	6	67.647	19.865	1.748
49	CD2	TRP	6	68.489	19.216	-0.249
50	NE1	TRP	6	66.801	20.249	0.758
51	CE2	TRP	6	67.268	19.885	-0.450
52	CE3	TRP	6	69.204	18.725	-1.329
53	CZ2	TRP	6	66.776	20.044	-1.737
54	CZ3	TRP	6	68.706	18.889	-2.614
55	CH2	TRP	6	67.497	19.547	-2.818
56	N	THR	7	68.832	15.952	4.241
57	CA	THR	7	68.347	14.565	4.241
58	C	THR	7	66.972	14.397	4.869
59	O	THR	7	66.620	13.279	5.267
60	CB	THR	7	69.338	13.662	4.968
61	OG1	THR	7	69.584	14.211	6.254
62	CG2	THR	7	70.667	13.559	4.230
63	N	TRP	8	66.145	15.426	4.815
64	CA	TRP	8	64.829	15.341	5.463
65	C	TRP	8	63.850	14.424	4.736
66	O	TRP	8	63.119	13.683	5.401
67	CB	TRP	8	64.236	16.736	5.575
68	CG	TRP	8	64.909	17.584	6.632
69	CD1	TRP	8	66.042	18.351	6.488
70	CD2	TRP	8	64.484	17.732	8.003
71	NE1	TRP	8	66.304	18.956	7.671
72	CE2	TRP	8	65.400	18.613	8.606
73	CE3	TRP	8	63.432	17.206	8.736
74	CZ2	TRP	8	65.248	18.955	9.941
75	CZ3	TRP	8	63.287	17.555	10.073
76	CH2	TRP	8	64.192	18.427	10.672
77	N	GLU	9	64.022	14.251	3.436

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
78	CA	GLU	9	63.165	13.288	2.744
79	C	GLU	9	63.736	11.876	2.825
80	O	GLU	9	62.951	10.924	2.810
81	CB	GLU	9	62.974	13.704	1.292
82	CG	GLU	9	62.138	12.703	0.494
83	CD	GLU	9	60.733	12.520	1.074
84	OE1	GLU	9	60.218	13.460	1.660
85	OE2	GLU	9	60.205	11.425	0.932
86	N	GLN	10	65.000	11.746	3.191
87	CA	GLN	10	65.587	10.415	3.314
88	C	GLN	10	65.265	9.845	4.689
89	O	GLN	10	64.872	8.675	4.803
90	CB	GLN	10	67.095	10.524	3.117
91	CG	GLN	10	67.767	9.156	3.158
92	CD	GLN	10	67.184	8.252	2.076
93	OE1	GLN	10	67.010	8.667	0.924
94	NE2	GLN	10	66.875	7.030	2.470
95	N	THR	11	65.152	10.741	5.655
96	CA	THR	11	64.696	10.337	6.985
97	C	THR	11	63.179	10.157	7.012
98	O	THR	11	62.702	9.226	7.671
99	CB	THR	11	65.136	11.374	8.015
100	OG1	THR	11	64.609	12.643	7.650
101	CG2	THR	11	66.655	11.493	8.075
102	N	PHE	12	62.475	10.827	6.111
103	CA	PHE	12	61.031	10.607	5.978
104	C	PHE	12	60.728	9.293	5.267
105	O	PHE	12	59.824	8.570	5.700
106	CB	PHE	12	60.423	11.759	5.188
107	CG	PHE	12	58.931	11.596	4.912
108	CD1	PHE	12	58.484	11.347	3.621
109	CD2	PHE	12	58.018	11.698	5.952
110	CE1	PHE	12	57.127	11.202	3.370
111	CE2	PHE	12	56.660	11.552	5.701
112	CZ	PHE	12	56.213	11.305	4.410
113	N	GLN	13	61.620	8.879	4.379
114	CA	GLN	13	61.487	7.576	3.728
115	C	GLN	13	61.769	6.453	4.707
116	O	GLN	13	60.977	5.509	4.772

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
117	CB	GLN	13	62.475	7.488	2.571
118	CG	GLN	13	61.995	8.282	1.364
119	CD	GLN	13	60.697	7.672	0.849
120	OE1	GLN	13	60.650	6.482	0.513
121	NE2	GLN	13	59.673	8.501	0.755
122	N	GLU	14	62.644	6.720	5.660
123	CA	GLU	14	62.944	5.739	6.698
124	C	GLU	14	61.857	5.680	7.765
125	O	GLU	14	61.523	4.570	8.190
126	CB	GLU	14	64.298	6.101	7.278
127	CG	GLU	14	65.345	5.855	6.200
128	CD	GLU	14	66.570	6.733	6.404
129	OE1	GLU	14	67.377	6.774	5.482
130	OE2	GLU	14	66.581	7.492	7.364
131	N	LEU	15	61.100	6.755	7.919
132	CA	LEU	15	59.904	6.708	8.771
133	C	LEU	15	58.801	5.883	8.114
134	O	LEU	15	58.217	5.016	8.777
135	CB	LEU	15	59.388	8.126	8.988
136	CG	LEU	15	60.333	8.961	9.841
137	CD1	LEU	15	59.917	10.427	9.831
138	CD2	LEU	15	60.395	8.424	11.266
139	N	ILE	16	58.750	5.931	6.792
140	CA	ILE	16	57.779	5.139	6.026
141	C	ILE	16	58.143	3.652	6.047
142	O	ILE	16	57.273	2.805	6.276
143	CB	ILE	16	57.809	5.630	4.577
144	CG1	ILE	16	57.439	7.105	4.455
145	CG2	ILE	16	56.908	4.785	3.685
146	CD1	ILE	16	56.014	7.388	4.912
147	N	GLN	17	59.438	3.380	6.075
148	CA	GLN	17	59.958	2.008	6.077
149	C	GLN	17	60.061	1.392	7.475
150	O	GLN	17	60.350	0.197	7.604
151	CB	GLN	17	61.339	2.077	5.435
152	CG	GLN	17	61.220	2.544	3.987
153	CD	GLN	17	62.514	3.193	3.497
154	OE1	GLN	17	63.462	3.411	4.265
155	NE2	GLN	17	62.463	3.657	2.260

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
156	N	GLU	18	59.846	2.196	8.504
157	CA	GLU	18	59.863	1.698	9.881
158	C	GLU	18	58.471	1.716	10.499
159	O	GLU	18	58.307	1.300	11.654
160	CB	GLU	18	60.790	2.573	10.715
161	CG	GLU	18	62.241	2.468	10.264
162	CD	GLU	18	63.088	3.487	11.020
163	OE1	GLU	18	63.617	3.123	12.061
164	OE2	GLU	18	63.227	4.600	10.529
165	N	ALA	19	57.511	2.261	9.769
166	CA	ALA	19	56.126	2.304	10.241
167	C	ALA	19	55.616	0.890	10.469
168	O	ALA	19	55.700	0.046	9.572
169	CB	ALA	19	55.283	2.994	9.179
170	N	LYS	20	55.062	0.674	11.652
171	CA	LYS	20	54.662	-0.667	12.120
172	C	LYS	20	55.750	-1.709	11.824
173	O	LYS	20	55.667	-2.437	10.825
174	CB	LYS	20	53.353	-1.045	11.433
175	CG	LYS	20	52.778	-2.356	11.964
176	CD	LYS	20	51.489	-2.735	11.240
177	CE	LYS	20	51.662	-3.947	10.325
178	NZ	LYS	20	52.642	-3.713	9.252
179	N	PRO	21	56.780	-1.744	12.657
180	CA	PRO	21	57.934	-2.596	12.376
181	C	PRO	21	57.607	-4.071	12.568
182	O	PRO	21	57.139	-4.495	13.632
183	CB	PRO	21	58.996	-2.153	13.336
184	CG	PRO	21	58.400	-1.152	14.313
185	CD	PRO	21	56.964	-0.942	13.870
186	N	ARG	22	57.814	-4.829	11.508
187	CA	ARG	22	57.665	-6.282	11.577
188	C	ARG	22	58.830	-6.888	12.348
189	O	ARG	22	59.915	-6.299	12.425
190	CB	ARG	22	57.618	-6.840	10.156
191	CG	ARG	22	56.477	-6.271	9.306
192	CD	ARG	22	55.117	-6.933	9.549
193	NE	ARG	22	54.500	-6.572	10.838
194	CZ	ARG	22	53.376	-7.126	11.294

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
195	NH1	ARG	22	52.712	-8.005	10.540
196	NH2	ARG	22	52.888	-6.762	12.482
197	N	ALA	23	58.583	-8.036	12.957
198	CA	ALA	23	59.638	-8.707	13.725
199	C	ALA	23	60.558	-9.504	12.806
200	O	ALA	23	61.741	-9.704	13.104
201	CB	ALA	23	58.989	-9.640	14.741
202	N	THR	24	60.010	-9.930	11.681
203	CA	THR	24	60.824	-10.551	10.635
204	C	THR	24	60.782	-9.670	9.399
205	O	THR	24	59.740	-9.074	9.092
206	CB	THR	24	60.278	-11.933	10.280
207	OG1	THR	24	59.027	-11.779	9.621
208	CG2	THR	24	60.084	-12.815	11.508
209	N	TRP	25	61.878	-9.650	8.660
210	CA	TRP	25	61.945	-8.893	7.401
211	C	TRP	25	61.341	-9.710	6.259
212	O	TRP	25	62.060	-10.302	5.447
213	CB	TRP	25	63.409	-8.588	7.092
214	CG	TRP	25	64.141	-7.840	8.191
215	CD1	TRP	25	65.006	-8.375	9.121
216	CD2	TRP	25	64.077	-6.422	8.459
217	NE1	TRP	25	65.443	-7.375	9.927
218	CE2	TRP	25	64.913	-6.192	9.567
219	CE3	TRP	25	63.395	-5.371	7.867
220	CZ2	TRP	25	65.051	-4.906	10.069
221	CZ3	TRP	25	63.540	-4.086	8.376
222	CH2	TRP	25	64.364	-3.855	9.473
223	N	THR	26	60.022	-9.800	6.257
224	CA	THR	26	59.317	-10.636	5.288
225	C	THR	26	58.136	-9.878	4.701
226	O	THR	26	57.999	-9.749	3.479
227	CB	THR	26	58.818	-11.874	6.025
228	OG1	THR	26	59.905	-12.426	6.755
229	CG2	THR	26	58.279	-12.931	5.067
230	N	LEU	27	57.279	-9.403	5.587
231	CA	LEU	27	56.130	-8.610	5.155
232	C	LEU	27	56.583	-7.166	4.993
233	O	LEU	27	57.350	-6.659	5.820

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
234	CB	LEU	27	55.024	-8.718	6.204
235	CG	LEU	27	53.717	-8.080	5.737
236	CD1	LEU	27	53.212	-8.740	4.458
237	CD2	LEU	27	52.650	-8.146	6.823
238	N	LYS	28	56.185	-6.559	3.890
239	CA	LYS	28	56.529	-5.164	3.621
240	C	LYS	28	55.845	-4.235	4.615
241	O	LYS	28	56.403	-3.902	5.667
242	CB	LYS	28	56.056	-4.826	2.213
243	CG	LYS	28	56.623	-5.804	1.193
244	CD	LYS	28	56.089	-5.511	-0.204
245	CE	LYS	28	56.629	-6.508	-1.223
246	NZ	LYS	28	56.116	-6.210	-2.570
247	N	LEU	29	54.615	-3.874	4.294
248	CA	LEU	29	53.884	-2.909	5.115
249	C	LEU	29	52.388	-2.954	4.823
250	O	LEU	29	51.957	-2.594	3.722
251	CB	LEU	29	54.434	-1.533	4.754
252	CG	LEU	29	53.779	-0.404	5.535
253	CD1	LEU	29	53.984	-0.583	7.034
254	CD2	LEU	29	54.339	0.933	5.070
255	N	ASP	30	51.608	-3.421	5.785
256	CA	ASP	30	50.150	-3.401	5.610
257	C	ASP	30	49.418	-3.513	6.944
258	O	ASP	30	49.411	-4.566	7.594
259	CB	ASP	30	49.714	-4.533	4.686
260	CG	ASP	30	48.526	-4.057	3.849
261	OD1	ASP	30	48.065	-2.957	4.124
262	OD2	ASP	30	48.305	-4.662	2.808
263	N	GLY	31	48.735	-2.435	7.293
264	CA	GLY	31	47.954	-2.384	8.534
265	C	GLY	31	46.479	-2.653	8.243
266	O	GLY	31	45.657	-1.732	8.186
267	N	ASN	32	46.175	-3.927	8.061
268	CA	ASN	32	44.821	-4.378	7.717
269	C	ASN	32	43.910	-4.231	8.934
270	O	ASN	32	44.071	-4.943	9.931
271	CB	ASN	32	44.917	-5.851	7.321
272	CG	ASN	32	46.188	-6.129	6.509

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
273	OD1	ASN	32	46.468	-5.482	5.490
274	ND2	ASN	32	46.946	-7.107	6.976
275	N	LEU	33	42.966	-3.308	8.850
276	CA	LEU	33	42.184	-2.942	10.035
277	C	LEU	33	40.677	-2.890	9.831
278	O	LEU	33	39.954	-3.879	10.008
279	CB	LEU	33	42.618	-1.552	10.493
280	CG	LEU	33	43.990	-1.548	11.155
281	CD1	LEU	33	44.451	-0.124	11.448
282	CD2	LEU	33	43.974	-2.385	12.429
283	N	GLN	34	40.228	-1.685	9.516
284	CA	GLN	34	38.813	-1.313	9.629
285	C	GLN	34	37.856	-1.973	8.638
286	O	GLN	34	38.208	-2.855	7.845
287	CB	GLN	34	38.693	0.198	9.512
288	CG	GLN	34	39.229	0.699	8.180
289	CD	GLN	34	38.834	2.156	8.000
290	OE1	GLN	34	39.191	3.021	8.811
291	NE2	GLN	34	38.066	2.397	6.954
292	N	LEU	35	36.617	-1.522	8.774
293	CA	LEU	35	35.435	-2.047	8.082
294	C	LEU	35	35.563	-2.205	6.568
295	O	LEU	35	36.282	-1.474	5.874
296	CB	LEU	35	34.288	-1.086	8.381
297	CG	LEU	35	33.026	-1.844	8.776
298	CD1	LEU	35	33.263	-2.665	10.040
299	CD2	LEU	35	31.855	-0.890	8.973
300	N	ASP	36	34.797	-3.176	6.091
301	CA	ASP	36	34.664	-3.540	4.669
302	C	ASP	36	34.007	-2.470	3.789
303	O	ASP	36	33.665	-1.373	4.254
304	CB	ASP	36	33.814	-4.813	4.613
305	CG	ASP	36	32.559	-4.654	5.475
306	OD1	ASP	36	31.758	-3.782	5.155
307	OD2	ASP	36	32.516	-5.278	6.523
308	N	CYS	37	33.598	-2.926	2.608
309	CA	CYS	37	32.997	-2.098	1.542
310	C	CYS	37	31.618	-1.483	1.825
311	O	CYS	37	31.107	-0.740	0.979

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
312	CB	CYS	37	32.857	-2.986	0.312
313	SG	CYS	37	34.390	-3.712	-0.306
314	N	LEU	38	31.049	-1.715	2.997
315	CA	LEU	38	29.815	-1.035	3.369
316	C	LEU	38	30.131	0.374	3.873
317	O	LEU	38	29.357	1.304	3.625
318	CB	LEU	38	29.135	-1.840	4.466
319	CG	LEU	38	27.778	-1.244	4.815
320	CD1	LEU	38	26.876	-1.203	3.585
321	CD2	LEU	38	27.113	-2.018	5.946
322	N	ALA	39	31.372	0.568	4.294
323	CA	ALA	39	31.862	1.889	4.706
324	C	ALA	39	32.523	2.653	3.555
325	O	ALA	39	33.080	3.737	3.770
326	CB	ALA	39	32.866	1.704	5.839
327	N	GLN	40	32.280	2.204	2.334
328	CA	GLN	40	32.992	2.713	1.155
329	C	GLN	40	32.498	4.078	0.656
330	O	GLN	40	33.200	4.746	-0.109
331	CB	GLN	40	32.791	1.661	0.078
332	CG	GLN	40	33.586	1.929	-1.184
333	CD	GLN	40	33.363	0.773	-2.143
334	OE1	GLN	40	34.130	0.585	-3.094
335	NE2	GLN	40	32.353	-0.026	-1.842
336	N	GLY	41	31.362	4.529	1.161
337	CA	GLY	41	30.885	5.881	0.853
338	C	GLY	41	30.988	6.764	2.093
339	O	GLY	41	30.665	7.959	2.068
340	N	TRP	42	31.442	6.158	3.176
341	CA	TRP	42	31.553	6.856	4.450
342	C	TRP	42	32.985	7.346	4.628
343	O	TRP	42	33.218	8.442	5.153
344	CB	TRP	42	31.147	5.886	5.564
345	CG	TRP	42	29.644	5.704	5.773
346	CD1	TRP	42	28.937	6.167	6.861
347	CD2	TRP	42	28.686	5.016	4.934
348	NE1	TRP	42	27.632	5.846	6.704
349	CE2	TRP	42	27.430	5.169	5.558
350	CE3	TRP	42	28.774	4.338	3.730

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
351	CZ2	TRP	42	26.289	4.666	4.953
352	CZ3	TRP	42	27.629	3.826	3.133
353	CH2	TRP	42	26.389	3.994	3.741
354	N	LYS	43	33.916	6.594	4.066
355	CA	LYS	43	35.318	7.015	4.037
356	C	LYS	43	35.593	7.820	2.775
357	O	LYS	43	34.869	7.687	1.782
358	CB	LYS	43	36.215	5.784	4.083
359	CG	LYS	43	35.958	4.989	5.357
360	CD	LYS	43	36.247	5.835	6.592
361	CE	LYS	43	35.760	5.161	7.868
362	NZ	LYS	43	35.897	6.075	9.011
363	N	GLN	44	36.535	8.741	2.872
364	CA	GLN	44	36.915	9.542	1.709
365	C	GLN	44	38.221	9.009	1.126
366	O	GLN	44	38.460	9.092	-0.085
367	CB	GLN	44	37.092	10.991	2.162
368	CG	GLN	44	37.173	11.963	0.989
369	CD	GLN	44	37.314	13.399	1.493
370	OE1	GLN	44	38.412	13.965	1.486
371	NE2	GLN	44	36.205	13.971	1.933
372	N	TYR	45	39.022	8.400	1.985
373	CA	TYR	45	40.313	7.856	1.543
374	C	TYR	45	40.349	6.345	1.722
375	O	TYR	45	40.511	5.846	2.841
376	CB	TYR	45	41.436	8.544	2.313
377	CG	TYR	45	41.444	10.055	2.087
378	CD1	TYR	45	41.956	10.577	0.906
379	CD2	TYR	45	40.921	10.904	3.054
380	CE1	TYR	45	41.935	11.948	0.687
381	CE2	TYR	45	40.895	12.273	2.833
382	CZ	TYR	45	41.398	12.792	1.649
383	OH	TYR	45	41.296	14.143	1.400
384	N	GLN	46	40.522	5.674	0.597
385	CA	GLN	46	40.290	4.227	0.504
386	C	GLN	46	41.478	3.387	0.974
387	O	GLN	46	41.362	2.179	1.213
388	CB	GLN	46	40.048	3.937	-0.974
389	CG	GLN	46	39.289	2.636	-1.187

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
390	CD	GLN	46	37.814	2.839	-0.850
391	OE1	GLN	46	37.442	3.178	0.281
392	NE2	GLN	46	36.987	2.637	-1.859
393	N	GLN	47	42.613	4.048	1.107
394	CA	GLN	47	43.861	3.409	1.514
395	C	GLN	47	44.146	3.627	3.000
396	O	GLN	47	45.066	3.014	3.556
397	CB	GLN	47	45.013	4.071	0.743
398	CG	GLN	47	44.707	4.423	-0.719
399	CD	GLN	47	44.300	5.898	-0.889
400	OE1	GLN	47	43.141	6.267	-0.650
401	NE2	GLN	47	45.243	6.714	-1.325
402	N	ARG	48	43.392	4.520	3.623
403	CA	ARG	48	43.770	4.996	4.959
404	C	ARG	48	42.718	4.719	6.030
405	O	ARG	48	41.578	5.189	5.954
406	CB	ARG	48	44.025	6.493	4.846
407	CG	ARG	48	44.994	6.758	3.701
408	CD	ARG	48	45.419	8.215	3.592
409	NE	ARG	48	46.252	8.406	2.392
410	CZ	ARG	48	47.566	8.163	2.331
411	NH1	ARG	48	48.237	7.790	3.424
412	NH2	ARG	48	48.221	8.348	1.184
413	N	ALA	49	43.140	3.976	7.040
414	CA	ALA	49	42.271	3.653	8.179
415	C	ALA	49	42.410	4.680	9.302
416	O	ALA	49	43.524	4.949	9.769
417	CB	ALA	49	42.641	2.269	8.697
418	N	PHE	50	41.277	5.164	9.784
419	CA	PHE	50	41.267	6.221	10.815
420	C	PHE	50	41.588	5.652	12.198
421	O	PHE	50	41.450	4.444	12.421
422	CB	PHE	50	39.916	6.942	10.822
423	CG	PHE	50	38.791	6.301	11.636
424	CD1	PHE	50	38.330	5.022	11.347
425	CD2	PHE	50	38.218	7.023	12.676
426	CE1	PHE	50	37.309	4.463	12.104
427	CE2	PHE	50	37.196	6.465	13.433
428	CZ	PHE	50	36.743	5.183	13.148

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
429	N	GLY	51	42.031	6.503	13.111
430	CA	GLY	51	42.378	5.984	14.438
431	C	GLY	51	42.724	7.024	15.502
432	O	GLY	51	43.130	8.164	15.241
433	N	TRP	52	42.565	6.576	16.734
434	CA	TRP	52	42.967	7.356	17.906
435	C	TRP	52	44.378	6.923	18.281
436	O	TRP	52	45.074	6.327	17.453
437	CB	TRP	52	42.052	7.023	19.083
438	CG	TRP	52	40.654	6.553	18.732
439	CD1	TRP	52	39.574	7.328	18.367
440	CD2	TRP	52	40.198	5.181	18.720
441	NE1	TRP	52	38.516	6.509	18.137
442	CE2	TRP	52	38.845	5.217	18.339
443	CE3	TRP	52	40.815	3.973	19.001
444	CZ2	TRP	52	38.129	4.035	18.235
445	CZ3	TRP	52	40.090	2.791	18.894
446	CH2	TRP	52	38.751	2.824	18.515
447	N	PHE	53	44.838	7.386	19.431
448	CA	PHE	53	46.009	6.787	20.101
449	C	PHE	53	46.247	7.356	21.491
450	O	PHE	53	46.292	8.572	21.718
451	CB	PHE	53	47.305	6.796	19.281
452	CG	PHE	53	47.734	8.075	18.571
453	CD1	PHE	53	48.018	8.010	17.215
454	CD2	PHE	53	47.876	9.274	19.253
455	CE1	PHE	53	48.417	9.150	16.535
456	CE2	PHE	53	48.277	10.415	18.572
457	CZ	PHE	53	48.545	10.353	17.213
458	N	ARG	54	46.360	6.433	22.427
459	CA	ARG	54	46.677	6.779	23.809
460	C	ARG	54	48.184	6.739	24.011
461	O	ARG	54	48.765	5.651	24.115
462	CB	ARG	54	46.051	5.727	24.715
463	CG	ARG	54	44.554	5.574	24.476
464	CD	ARG	54	43.987	4.431	25.311
465	NE	ARG	54	44.295	4.620	26.737
466	CZ	ARG	54	44.938	3.711	27.475
467	NH1	ARG	54	45.265	3.989	28.739

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
468	NH2	ARG	54	45.324	2.557	26.925
469	N	CYS	55	48.816	7.898	24.056
470	CA	CYS	55	50.252	7.900	24.326
471	C	CYS	55	50.474	7.725	25.823
472	O	CYS	55	50.337	8.672	26.606
473	CB	CYS	55	50.904	9.189	23.834
474	SG	CYS	55	52.645	9.318	24.300
475	N	SER	56	50.952	6.550	26.197
476	CA	SER	56	51.107	6.216	27.620
477	C	SER	56	52.368	6.807	28.251
478	O	SER	56	52.391	7.041	29.464
479	CB	SER	56	51.114	4.700	27.753
480	OG	SER	56	49.883	4.225	27.221
481	N	SER	57	53.295	7.256	27.420
482	CA	SER	57	54.468	7.980	27.929
483	C	SER	57	54.223	9.491	27.956
484	O	SER	57	55.143	10.265	28.237
485	CB	SER	57	55.676	7.683	27.047
486	OG	SER	57	55.453	8.285	25.779
487	N	CYS	58	53.022	9.895	27.575
488	CA	CYS	58	52.652	11.305	27.555
489	C	CYS	58	51.558	11.554	28.582
490	O	CYS	58	51.502	12.616	29.212
491	CB	CYS	58	52.034	11.622	26.197
492	SG	CYS	58	52.950	11.272	24.680
493	N	GLN	59	50.694	10.553	28.690
494	CA	GLN	59	49.461	10.571	29.491
495	C	GLN	59	48.412	11.451	28.815
496	O	GLN	59	47.793	12.316	29.444
497	CB	GLN	59	49.730	10.996	30.932
498	CG	GLN	59	50.522	9.925	31.678
499	CD	GLN	59	49.690	8.651	31.811
500	OE1	GLN	59	48.530	8.694	32.234
501	NE2	GLN	59	50.270	7.534	31.405
502	N	ARG	60	48.204	11.177	27.535
503	CA	ARG	60	47.206	11.908	26.738
504	C	ARG	60	46.750	11.087	25.531
505	O	ARG	60	47.564	10.518	24.791
506	CB	ARG	60	47.804	13.230	26.264

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
507	CG	ARG	60	46.810	14.051	25.447
508	CD	ARG	60	47.404	15.372	24.979
509	NE	ARG	60	47.664	16.273	26.110
510	CZ	ARG	60	48.505	17.307	26.037
511	NH1	ARG	60	49.214	17.512	24.925
512	NH2	ARG	60	48.677	18.102	27.095
513	N	SER	61	45.441	10.961	25.399
514	CA	SER	61	44.861	10.283	24.240
515	C	SER	61	44.381	11.289	23.195
516	O	SER	61	43.637	12.231	23.501
517	CB	SER	61	43.694	9.432	24.716
518	OG	SER	61	44.193	8.548	25.709
519	N	TRP	62	44.829	11.076	21.971
520	CA	TRP	62	44.413	11.902	20.833
521	C	TRP	62	43.474	11.110	19.922
522	O	TRP	62	43.346	9.886	20.071
523	CB	TRP	62	45.635	12.283	20.012
524	CG	TRP	62	46.753	13.053	20.694
525	CD1	TRP	62	47.896	12.528	21.251
526	CD2	TRP	62	46.843	14.488	20.852
527	NE1	TRP	62	48.679	13.550	21.683
528	CE2	TRP	62	48.092	14.734	21.449
529	CE3	TRP	62	46.019	15.537	20.489
530	CZ2	TRP	62	48.503	16.045	21.659
531	CZ3	TRP	62	46.434	16.845	20.714
532	CH2	TRP	62	47.672	17.096	21.296
533	N	ALA	63	42.806	11.804	19.014
534	CA	ALA	63	41.960	11.117	18.022
535	C	ALA	63	41.979	11.780	16.639
536	O	ALA	63	41.362	12.839	16.451
537	CB	ALA	63	40.531	11.086	18.552
538	N	SER	64	42.599	11.115	15.673
539	CA	SER	64	42.652	11.628	14.291
540	C	SER	64	41.540	11.025	13.429
541	O	SER	64	41.436	9.799	13.272
542	CB	SER	64	44.003	11.274	13.678
543	OG	SER	64	45.022	11.910	14.437
544	N	ALA	65	40.771	11.895	12.796
545	CA	ALA	65	39.628	11.418	12.003

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
546	C	ALA	65	39.378	12.214	10.722
547	O	ALA	65	39.268	13.445	10.753
548	CB	ALA	65	38.383	11.466	12.879
549	N	GLN	66	39.141	11.480	9.646
550	CA	GLN	66	38.849	12.064	8.325
551	C	GLN	66	37.575	12.902	8.309
552	O	GLN	66	36.730	12.797	9.207
553	CB	GLN	66	38.661	10.934	7.322
554	CG	GLN	66	37.552	10.003	7.789
555	CD	GLN	66	36.793	9.446	6.596
556	OE1	GLN	66	37.390	8.973	5.617
557	NE2	GLN	66	35.493	9.671	6.632
558	N	VAL	67	37.452	13.726	7.280
559	CA	VAL	67	36.264	14.572	7.096
560	C	VAL	67	35.017	13.780	6.697
561	O	VAL	67	34.501	12.973	7.488
562	CB	VAL	67	36.582	15.636	6.045
563	CG1	VAL	67	36.968	16.963	6.689
564	CG2	VAL	67	37.655	15.168	5.063
565	N	GLN	68	34.534	14.038	5.487
566	CA	GLN	68	33.296	13.443	4.953
567	C	GLN	68	32.028	13.947	5.617
568	O	GLN	68	32.035	14.668	6.623
569	CB	GLN	68	33.313	11.924	5.059
570	CG	GLN	68	34.148	11.295	3.962
571	CD	GLN	68	33.424	11.449	2.638
572	OE1	GLN	68	33.669	12.410	1.897
573	NE2	GLN	68	32.479	10.555	2.413
574	N	ILE	69	30.937	13.410	5.096
575	CA	ILE	69	29.565	13.718	5.517
576	C	ILE	69	29.112	13.036	6.815
577	O	ILE	69	27.901	12.964	7.066
578	CB	ILE	69	28.658	13.261	4.382
579	CG1	ILE	69	28.864	11.773	4.114
580	CG2	ILE	69	28.933	14.069	3.120
581	CD1	ILE	69	27.980	11.285	2.973
582	N	LEU	70	30.040	12.681	7.694
583	CA	LEU	70	29.719	11.950	8.933
584	C	LEU	70	29.414	12.890	10.100

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
585	O	LEU	70	29.636	12.538	11.266
586	CB	LEU	70	30.895	11.051	9.294
587	CG	LEU	70	31.123	9.997	8.219
588	CD1	LEU	70	32.288	9.085	8.586
589	CD2	LEU	70	29.853	9.183	7.999
590	N	CYS	71	28.618	13.902	9.792
591	CA	CYS	71	28.340	15.018	10.695
592	C	CYS	71	27.328	14.646	11.774
593	O	CYS	71	27.293	15.275	12.837
594	CB	CYS	71	27.760	16.129	9.828
595	SG	CYS	71	28.586	16.325	8.231
596	N	HIS	72	26.643	13.533	11.574
597	CA	HIS	72	25.706	13.054	12.584
598	C	HIS	72	26.414	12.267	13.683
599	O	HIS	72	26.089	12.430	14.865
600	CB	HIS	72	24.665	12.160	11.918
601	CG	HIS	72	23.684	12.866	10.998
602	ND1	HIS	72	23.332	14.166	11.028
603	CD2	HIS	72	22.980	12.288	9.969
604	CE1	HIS	72	22.429	14.409	10.058
605	NE2	HIS	72	22.211	13.247	9.403
606	N	THR	73	27.525	11.650	13.318
607	CA	THR	73	28.272	10.805	14.249
608	C	THR	73	29.441	11.566	14.868
609	O	THR	73	29.971	11.156	15.910
610	CB	THR	73	28.752	9.588	13.463
611	OG1	THR	73	27.592	8.922	12.984
612	CG2	THR	73	29.536	8.597	14.317
613	N	TYR	74	29.636	12.792	14.407
614	CA	TYR	74	30.763	13.606	14.872
615	C	TYR	74	30.546	14.291	16.225
616	O	TYR	74	31.431	15.032	16.664
617	CB	TYR	74	31.104	14.634	13.802
618	CG	TYR	74	32.507	14.444	13.229
619	CD1	TYR	74	33.203	15.531	12.716
620	CD2	TYR	74	33.086	13.181	13.221
621	CE1	TYR	74	34.484	15.357	12.206
622	CE2	TYR	74	34.366	13.005	12.711
623	CZ	TYR	74	35.064	14.095	12.209

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
624	OH	TYR	74	36.359	13.934	11.765
625	N	TRP	75	29.392	14.096	16.849
626	CA	TRP	75	29.211	14.484	18.257
627	C	TRP	75	28.880	13.263	19.117
628	O	TRP	75	28.809	13.342	20.348
629	CB	TRP	75	28.103	15.524	18.362
630	CG	TRP	75	28.417	16.809	17.618
631	CD1	TRP	75	27.630	17.421	16.674
632	CD2	TRP	75	29.601	17.625	17.766
633	NE1	TRP	75	28.278	18.523	16.229
634	CE2	TRP	75	29.466	18.677	16.835
635	CE3	TRP	75	30.738	17.535	18.553
636	CZ2	TRP	75	30.491	19.600	16.692
637	CZ3	TRP	75	31.752	18.474	18.410
638	CH2	TRP	75	31.630	19.501	17.481
639	N	GLU	76	28.813	12.118	18.457
640	CA	GLU	76	28.383	10.862	19.081
641	C	GLU	76	29.568	10.022	19.573
642	O	GLU	76	29.371	8.994	20.234
643	CB	GLU	76	27.595	10.105	18.014
644	CG	GLU	76	26.942	8.825	18.519
645	CD	GLU	76	26.329	8.083	17.338
646	OE1	GLU	76	26.777	8.322	16.223
647	OE2	GLU	76	25.384	7.339	17.559
648	N	HIS	77	30.767	10.561	19.419
649	CA	HIS	77	32.024	9.851	19.718
650	C	HIS	77	32.433	9.804	21.195
651	O	HIS	77	33.633	9.648	21.454
652	CB	HIS	77	33.141	10.560	18.950
653	CG	HIS	77	33.315	12.042	19.264
654	ND1	HIS	77	33.834	12.583	20.385
655	CD2	HIS	77	32.983	13.091	18.440
656	CE1	HIS	77	33.820	13.927	20.287
657	NE2	HIS	77	33.293	14.241	19.083
658	N	TRP	78	31.492	9.942	22.120
659	CA	TRP	78	31.793	10.171	23.544
660	C	TRP	78	32.500	11.513	23.698
661	O	TRP	78	33.657	11.689	23.289
662	CB	TRP	78	32.617	9.032	24.144

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
663	CG	TRP	78	31.887	7.703	24.176
664	CD1	TRP	78	30.939	7.318	25.097
665	CD2	TRP	78	32.038	6.598	23.256
666	NE1	TRP	78	30.515	6.068	24.782
667	CE2	TRP	78	31.143	5.603	23.686
668	CE3	TRP	78	32.830	6.390	22.137
669	CZ2	TRP	78	31.048	4.411	22.983
670	CZ3	TRP	78	32.729	5.193	21.438
671	CH2	TRP	78	31.842	4.207	21.860
672	N	THR	79	31.769	12.445	24.286
673	CA	THR	79	32.172	13.856	24.350
674	C	THR	79	33.588	14.108	24.862
675	O	THR	79	33.996	13.703	25.959
676	CB	THR	79	31.170	14.607	25.220
677	OG1	THR	79	31.724	15.887	25.502
678	CG2	THR	79	30.925	13.904	26.551
679	N	SER	80	34.326	14.784	24.003
680	CA	SER	80	35.664	15.262	24.323
681	C	SER	80	35.531	16.655	24.912
682	O	SER	80	34.447	17.018	25.387
683	CB	SER	80	36.464	15.296	23.030
684	OG	SER	80	36.374	13.992	22.471
685	N	GLN	81	36.628	17.376	25.023
686	CA	GLN	81	36.478	18.772	25.427
687	C	GLN	81	36.656	19.679	24.218
688	O	GLN	81	36.126	20.799	24.180
689	CB	GLN	81	37.490	19.117	26.512
690	CG	GLN	81	36.846	19.870	27.677
691	CD	GLN	81	36.240	21.189	27.208
692	OE1	GLN	81	36.967	22.089	26.768
693	NE2	GLN	81	34.923	21.275	27.290
694	N	GLY	82	37.280	19.151	23.182
695	CA	GLY	82	37.508	19.995	22.013
696	C	GLY	82	37.600	19.269	20.681
697	O	GLY	82	37.780	18.045	20.591
698	N	GLN	83	37.438	20.092	19.660
699	CA	GLN	83	37.580	19.702	18.256
700	C	GLN	83	38.483	20.729	17.579
701	O	GLN	83	38.047	21.850	17.282

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
702	CB	GLN	83	36.199	19.713	17.607
703	CG	GLN	83	36.247	19.430	16.105
704	CD	GLN	83	36.717	18.005	15.820
705	OE1	GLN	83	37.907	17.757	15.590
706	NE2	GLN	83	35.760	17.093	15.797
707	N	VAL	84	39.745	20.381	17.418
708	CA	VAL	84	40.702	21.348	16.873
709	C	VAL	84	41.019	21.043	15.414
710	O	VAL	84	41.603	20.004	15.080
711	CB	VAL	84	41.971	21.319	17.716
712	CG1	VAL	84	43.013	22.307	17.201
713	CG2	VAL	84	41.663	21.584	19.186
714	N	ARG	85	40.611	21.957	14.554
715	CA	ARG	85	40.841	21.809	13.119
716	C	ARG	85	41.673	22.956	12.540
717	O	ARG	85	41.781	24.030	13.134
718	CB	ARG	85	39.485	21.730	12.436
719	CG	ARG	85	38.830	20.379	12.704
720	CD	ARG	85	37.553	20.190	11.892
721	NE	ARG	85	36.400	20.891	12.477
722	CZ	ARG	85	35.175	20.356	12.459
723	NH1	ARG	85	34.974	19.165	11.889
724	NH2	ARG	85	34.149	21.011	13.005
725	N	MET	86	42.355	22.650	11.449
726	CA	MET	86	43.134	23.637	10.678
727	C	MET	86	42.809	23.455	9.197
728	O	MET	86	43.698	23.105	8.410
729	CB	MET	86	44.629	23.390	10.880
730	CG	MET	86	45.051	23.335	12.346
731	SD	MET	86	46.828	23.484	12.637
732	CE	MET	86	47.419	22.118	11.615
733	N	ARG	87	41.586	23.793	8.815
734	CA	ARG	87	41.036	23.306	7.545
735	C	ARG	87	40.893	24.320	6.416
736	O	ARG	87	40.509	25.482	6.599
737	CB	ARG	87	39.645	22.734	7.815
738	CG	ARG	87	39.614	21.726	8.963
739	CD	ARG	87	39.949	20.283	8.585
740	NE	ARG	87	41.248	20.150	7.910

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
741	CZ	ARG	87	42.427	19.932	8.494
742	NH1	ARG	87	43.533	19.873	7.752
743	NH2	ARG	87	42.505	19.791	9.816
744	N	LEU	88	41.161	23.807	5.231
745	CA	LEU	88	40.748	24.478	3.999
746	C	LEU	88	39.349	23.970	3.644
747	O	LEU	88	39.185	22.783	3.332
748	CB	LEU	88	41.736	24.153	2.886
749	CG	LEU	88	43.143	24.636	3.225
750	CD1	LEU	88	44.134	24.199	2.153
751	CD2	LEU	88	43.184	26.151	3.405
752	N	PHE	89	38.450	24.930	3.474
753	CA	PHE	89	36.976	24.761	3.385
754	C	PHE	89	36.428	23.607	2.540
755	O	PHE	89	36.787	22.431	2.705
756	CB	PHE	89	36.399	26.038	2.780
757	CG	PHE	89	36.702	27.328	3.531
758	CD1	PHE	89	37.623	28.230	3.011
759	CD2	PHE	89	36.038	27.616	4.715
760	CE1	PHE	89	37.891	29.415	3.688
761	CE2	PHE	89	36.304	28.797	5.389
762	CZ	PHE	89	37.229	29.697	4.878
763	N	GLY	90	35.554	23.975	1.611
764	CA	GLY	90	34.838	23.007	0.756
765	C	GLY	90	35.645	22.446	-0.417
766	O	GLY	90	35.196	22.435	-1.567
767	N	GLN	91	36.833	21.972	-0.090
768	CA	GLN	91	37.701	21.260	-1.009
769	C	GLN	91	37.708	19.809	-0.556
770	O	GLN	91	37.647	18.878	-1.367
771	CB	GLN	91	39.093	21.857	-0.888
772	CG	GLN	91	40.080	21.192	-1.839
773	CD	GLN	91	41.477	21.748	-1.606
774	OE1	GLN	91	42.080	22.348	-2.501
775	NE2	GLN	91	41.962	21.569	-0.389
776	N	ARG	92	37.723	19.637	0.758
777	CA	ARG	92	37.655	18.295	1.348
778	C	ARG	92	36.572	18.280	2.414
779	O	ARG	92	35.880	17.274	2.639

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
780	CB	ARG	92	38.975	17.981	2.045
781	CG	ARG	92	40.198	18.197	1.164
782	CD	ARG	92	40.328	17.174	0.044
783	NE	ARG	92	41.560	17.443	-0.713
784	CZ	ARG	92	41.583	17.638	-2.033
785	NH1	ARG	92	42.735	17.927	-2.642
786	NH2	ARG	92	40.451	17.579	-2.737
787	N	CYS	93	36.439	19.427	3.057
788	CA	CYS	93	35.506	19.575	4.167
789	C	CYS	93	34.214	20.270	3.759
790	O	CYS	93	34.144	21.499	3.623
791	CB	CYS	93	36.202	20.388	5.252
792	SG	CYS	93	35.238	20.706	6.749
793	N	GLN	94	33.188	19.458	3.582
794	CA	GLN	94	31.827	19.977	3.457
795	C	GLN	94	31.468	20.599	4.804
796	O	GLN	94	32.029	20.177	5.821
797	CB	GLN	94	30.897	18.793	3.189
798	CG	GLN	94	29.451	19.215	2.939
799	CD	GLN	94	28.511	18.058	3.248
800	OE1	GLN	94	28.299	17.171	2.414
801	NE2	GLN	94	27.971	18.079	4.455
802	N	LYS	95	30.664	21.649	4.807
803	CA	LYS	95	30.139	22.181	6.075
804	C	LYS	95	29.480	21.028	6.830
805	O	LYS	95	28.605	20.328	6.300
806	CB	LYS	95	29.131	23.305	5.818
807	CG	LYS	95	29.761	24.698	5.686
808	CD	LYS	95	30.614	24.889	4.433
809	CE	LYS	95	31.182	26.300	4.341
810	NZ	LYS	95	30.105	27.296	4.218
811	N	CYS	96	29.968	20.787	8.035
812	CA	CYS	96	29.652	19.526	8.704
813	C	CYS	96	29.704	19.627	10.219
814	O	CYS	96	30.722	20.016	10.804
815	CB	CYS	96	30.697	18.523	8.231
816	SG	CYS	96	30.499	16.781	8.671
817	N	SER	97	28.539	19.387	10.804
818	CA	SER	97	28.279	19.210	12.248

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
819	C	SER	97	28.364	20.467	13.121
820	O	SER	97	27.664	20.512	14.136
821	CB	SER	97	29.189	18.122	12.829
822	OG	SER	97	30.450	18.676	13.184
823	N	TRP	98	29.144	21.467	12.755
824	CA	TRP	98	29.154	22.682	13.563
825	C	TRP	98	28.168	23.688	12.991
826	O	TRP	98	28.102	23.910	11.774
827	CB	TRP	98	30.568	23.258	13.666
828	CG	TRP	98	30.978	24.282	12.624
829	CD1	TRP	98	30.847	25.648	12.737
830	CD2	TRP	98	31.596	24.036	11.340
831	NE1	TRP	98	31.326	26.224	11.607
832	CE2	TRP	98	31.782	25.297	10.745
833	CE3	TRP	98	31.989	22.885	10.676
834	CZ2	TRP	98	32.351	25.385	9.482
835	CZ3	TRP	98	32.563	22.982	9.414
836	CH2	TRP	98	32.741	24.227	8.820
837	N	SER	99	27.308	24.182	13.864
838	CA	SER	99	26.377	25.240	13.478
839	C	SER	99	27.171	26.499	13.188
840	O	SER	99	28.078	26.843	13.953
841	CB	SER	99	25.425	25.487	14.635
842	OG	SER	99	24.567	26.560	14.279
843	N	GLN	100	26.870	27.164	12.086
844	CA	GLN	100	27.653	28.350	11.725
845	C	GLN	100	27.033	29.636	12.254
846	O	GLN	100	27.720	30.655	12.392
847	CB	GLN	100	27.744	28.438	10.211
848	CG	GLN	100	28.372	27.186	9.618
849	CD	GLN	100	28.526	27.378	8.118
850	OE1	GLN	100	27.674	26.952	7.330
851	NE2	GLN	100	29.582	28.080	7.747
852	N	TYR	101	25.758	29.578	12.585
853	CA	TYR	101	25.094	30.767	13.110
854	C	TYR	101	24.384	30.429	14.408
855	O	TYR	101	23.904	29.305	14.595
856	CB	TYR	101	24.099	31.292	12.082
857	CG	TYR	101	24.679	31.558	10.696

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
858	CD1	TYR	101	24.334	30.734	9.630
859	CD2	TYR	101	25.540	32.630	10.495
860	CE1	TYR	101	24.865	30.970	8.369
861	CE2	TYR	101	26.072	32.866	9.234
862	CZ	TYR	101	25.736	32.033	8.176
863	OH	TYR	101	26.284	32.250	6.931
864	N	GLU	102	24.210	31.447	15.232
865	CA	GLU	102	23.624	31.271	16.566
866	C	GLU	102	22.137	30.928	16.498
867	O	GLU	102	21.646	30.119	17.295
868	CB	GLU	102	23.791	32.563	17.380
869	CG	GLU	102	25.195	32.823	17.948
870	CD	GLU	102	26.152	33.471	16.945
871	OE1	GLU	102	25.669	33.866	15.889
872	OE2	GLU	102	27.325	33.646	17.257
873	N	MET	103	21.467	31.411	15.465
874	CA	MET	103	20.043	31.105	15.303
875	C	MET	103	19.807	29.722	14.664
876	O	MET	103	19.193	28.897	15.349
877	CB	MET	103	19.348	32.233	14.543
878	CG	MET	103	19.498	33.565	15.270
879	SD	MET	103	18.749	34.985	14.441
880	CE	MET	103	17.046	34.386	14.346
881	N	PRO	104	20.223	29.441	13.428
882	CA	PRO	104	20.068	28.078	12.899
883	C	PRO	104	21.112	27.099	13.439
884	O	PRO	104	22.144	26.850	12.800
885	CB	PRO	104	20.214	28.208	11.416
886	CG	PRO	104	20.788	29.577	11.098
887	CD	PRO	104	20.861	30.309	12.427
888	N	GLU	105	20.808	26.529	14.593
889	CA	GLU	105	21.616	25.446	15.158
890	C	GLU	105	21.392	24.181	14.333
891	O	GLU	105	20.267	23.912	13.895
892	CB	GLU	105	21.241	25.217	16.629
893	CG	GLU	105	20.088	24.240	16.898
894	CD	GLU	105	18.711	24.794	16.521
895	OE1	GLU	105	18.573	26.010	16.510
896	OE2	GLU	105	17.806	23.993	16.343

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
897	N	PHE	106	22.450	23.432	14.094
898	CA	PHE	106	22.298	22.241	13.254
899	C	PHE	106	21.886	21.024	14.069
900	O	PHE	106	22.441	20.739	15.140
901	CB	PHE	106	23.584	21.969	12.480
902	CG	PHE	106	23.838	22.944	11.330
903	CD1	PHE	106	25.131	23.150	10.868
904	CD2	PHE	106	22.774	23.605	10.728
905	CE1	PHE	106	25.364	24.035	9.823
906	CE2	PHE	106	23.006	24.491	9.684
907	CZ	PHE	106	24.302	24.708	9.233
908	N	SER	107	20.835	20.375	13.601
909	CA	SER	107	20.364	19.152	14.250
910	C	SER	107	20.829	17.903	13.508
911	O	SER	107	20.609	17.715	12.304
912	CB	SER	107	18.845	19.164	14.339
913	OG	SER	107	18.459	17.961	14.992
914	N	SER	108	21.418	17.018	14.284
915	CA	SER	108	21.891	15.729	13.797
916	C	SER	108	20.815	14.686	14.063
917	O	SER	108	20.882	13.911	15.028
918	CB	SER	108	23.172	15.393	14.539
919	OG	SER	108	23.510	14.051	14.247
920	N	ASP	109	19.943	14.542	13.081
921	CA	ASP	109	18.690	13.807	13.287
922	C	ASP	109	18.847	12.298	13.453
923	O	ASP	109	18.188	11.748	14.342
924	CB	ASP	109	17.771	14.088	12.102
925	CG	ASP	109	17.431	15.576	12.023
926	OD1	ASP	109	17.212	16.048	10.918
927	OD2	ASP	109	17.358	16.201	13.073
928	N	SER	110	19.892	11.713	12.890
929	CA	SER	110	20.036	10.252	12.987
930	C	SER	110	20.690	9.793	14.292
931	O	SER	110	20.758	8.587	14.546
932	CB	SER	110	20.885	9.747	11.829
933	OG	SER	110	22.231	10.106	12.102
934	N	THR	111	21.200	10.726	15.082
935	CA	THR	111	21.799	10.367	16.371

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
936	C	THR	111	21.166	11.176	17.497
937	O	THR	111	21.674	11.171	18.626
938	CB	THR	111	23.297	10.639	16.333
939	OG1	THR	111	23.471	12.026	16.104
940	CG2	THR	111	23.991	9.875	15.211
941	N	MET	112	20.122	11.915	17.146
942	CA	MET	112	19.394	12.806	18.064
943	C	MET	112	20.299	13.848	18.720
944	O	MET	112	20.246	14.056	19.937
945	CB	MET	112	18.704	11.966	19.132
946	CG	MET	112	17.700	11.000	18.512
947	SD	MET	112	16.873	9.895	19.678
948	CE	MET	112	18.329	9.045	20.333
949	N	ARG	113	21.105	14.512	17.908
950	CA	ARG	113	21.982	15.568	18.427
951	C	ARG	113	21.484	16.945	18.004
952	O	ARG	113	20.790	17.096	16.993
953	CB	ARG	113	23.409	15.375	17.920
954	CG	ARG	113	24.041	14.068	18.383
955	CD	ARG	113	24.183	14.023	19.896
956	NE	ARG	113	24.822	12.777	20.340
957	CZ	ARG	113	25.563	12.712	21.447
958	NH1	ARG	113	25.780	13.816	22.167
959	NH2	ARG	113	26.110	11.553	21.820
960	N	ILE	114	21.787	17.932	18.826
961	CA	ILE	114	21.487	19.337	18.519
962	C	ILE	114	22.691	20.196	18.899
963	O	ILE	114	23.081	20.226	20.073
964	CB	ILE	114	20.264	19.782	19.326
965	CG1	ILE	114	19.005	19.019	18.925
966	CG2	ILE	114	20.030	21.282	19.187
967	CD1	ILE	114	17.799	19.471	19.741
968	N	LEU	115	23.313	20.827	17.916
969	CA	LEU	115	24.472	21.675	18.210
970	C	LEU	115	24.177	23.158	18.005
971	O	LEU	115	24.090	23.647	16.869
972	CB	LEU	115	25.636	21.313	17.306
973	CG	LEU	115	26.878	22.015	17.834
974	CD1	LEU	115	27.394	21.286	19.064

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
975	CD2	LEU	115	27.967	22.097	16.785
976	N	SER	116	24.100	23.872	19.113
977	CA	SER	116	23.919	25.326	19.070
978	C	SER	116	25.248	26.060	18.925
979	O	SER	116	26.234	25.738	19.602
980	CB	SER	116	23.242	25.779	20.357
981	OG	SER	116	23.206	27.201	20.345
982	N	ASN	117	25.279	26.993	17.991
983	CA	ASN	117	26.425	27.893	17.856
984	C	ASN	117	26.297	28.989	18.903
985	O	ASN	117	25.397	29.834	18.821
986	CB	ASN	117	26.369	28.483	16.456
987	CG	ASN	117	27.415	29.565	16.231
988	OD1	ASN	117	27.175	30.742	16.533
989	ND2	ASN	117	28.444	29.195	15.495
990	N	LEU	118	27.166	28.969	19.895
991	CA	LEU	118	26.982	29.901	21.009
992	C	LEU	118	27.676	31.234	20.744
993	O	LEU	118	27.018	32.200	20.344
994	CB	LEU	118	27.491	29.245	22.282
995	CG	LEU	118	26.501	29.415	23.427
996	CD1	LEU	118	25.131	28.872	23.034
997	CD2	LEU	118	27.010	28.732	24.693
998	N	VAL	119	28.963	31.315	21.030
999	CA	VAL	119	29.709	32.536	20.726
1000	C	VAL	119	30.913	32.181	19.867
1001	O	VAL	119	31.441	31.061	19.940
1002	CB	VAL	119	30.139	33.236	22.013
1003	CG1	VAL	119	28.946	33.830	22.754
1004	CG2	VAL	119	30.941	32.317	22.928
1005	N	GLN	120	31.259	33.096	18.980
1006	CA	GLN	120	32.352	32.844	18.041
1007	C	GLN	120	33.471	33.872	18.133
1008	O	GLN	120	33.252	35.072	17.925
1009	CB	GLN	120	31.786	32.901	16.629
1010	CG	GLN	120	30.542	32.037	16.501
1011	CD	GLN	120	30.068	31.997	15.057
1012	OE1	GLN	120	30.556	31.173	14.275
1013	NE2	GLN	120	28.993	32.713	14.787

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1014	N	HIS	121	34.670	33.388	18.403
1015	CA	HIS	121	35.856	34.232	18.238
1016	C	HIS	121	36.186	34.284	16.759
1017	O	HIS	121	36.560	33.266	16.157
1018	CB	HIS	121	37.065	33.665	18.974
1019	CG	HIS	121	37.064	33.787	20.481
1020	ND1	HIS	121	38.040	33.350	21.299
1021	CD2	HIS	121	36.102	34.364	21.277
1022	CE1	HIS	121	37.710	33.628	22.575
1023	NE2	HIS	121	36.511	34.255	22.561
1024	N	ILE	122	36.004	35.464	16.194
1025	CA	ILE	122	36.266	35.722	14.774
1026	C	ILE	122	37.738	35.445	14.466
1027	O	ILE	122	38.571	35.484	15.378
1028	CB	ILE	122	35.892	37.187	14.523
1029	CG1	ILE	122	34.453	37.440	14.966
1030	CG2	ILE	122	36.046	37.599	13.060
1031	CD1	ILE	122	33.451	36.668	14.111
1032	N	LEU	123	37.984	34.947	13.264
1033	CA	LEU	123	39.346	34.675	12.793
1034	C	LEU	123	40.246	35.871	13.054
1035	O	LEU	123	39.970	36.989	12.602
1036	CB	LEU	123	39.287	34.432	11.290
1037	CG	LEU	123	38.357	33.278	10.929
1038	CD1	LEU	123	37.956	33.344	9.461
1039	CD2	LEU	123	38.973	31.925	11.270
1040	N	LYS	124	41.285	35.629	13.835
1041	CA	LYS	124	42.246	36.677	14.177
1042	C	LYS	124	42.990	37.053	12.907
1043	O	LYS	124	42.681	38.049	12.246
1044	CB	LYS	124	43.230	36.128	15.208
1045	CG	LYS	124	42.521	35.400	16.349
1046	CD	LYS	124	41.620	36.309	17.182
1047	CE	LYS	124	40.817	35.486	18.183
1048	NZ	LYS	124	39.936	36.337	18.998
1049	N	LYS	125	43.981	36.245	12.583
1050	CA	LYS	125	44.620	36.369	11.275
1051	C	LYS	125	44.498	35.062	10.501
1052	O	LYS	125	44.380	35.061	9.270

Atom No.	Atom Type	Résidue	Residue Position	X Coord.	Y Coord.	Z Coord.
1053	CB	LYS	125	46.088	36.737	11.469
1054	CG	LYS	125	46.842	37.025	10.165
1055	CD	LYS	125	46.495	38.373	9.523
1056	CE	LYS	125	45.202	38.363	8.707
1057	NZ	LYS	125	44.914	39.696	8.156
1058	N	TYR	126	44.433	33.964	11.237
1059	CA	TYR	126	44.397	32.647	10.594
1060	C	TYR	126	43.604	31.610	11.381
1061	O	TYR	126	43.540	30.445	10.964
1062	CB	TYR	126	45.834	32.157	10.403
1063	CG	TYR	126	46.649	31.891	11.676
1064	CD1	TYR	126	47.363	32.917	12.286
1065	CD2	TYR	126	46.710	30.604	12.196
1066	CE1	TYR	126	48.099	32.667	13.436
1067	CE2	TYR	126	47.448	30.351	13.345
1068	CZ	TYR	126	48.134	31.385	13.966
1069	OH	TYR	126	48.828	31.143	15.131
1070	N	TYR	127	43.041	32.009	12.510
1071	CA	TYR	127	42.316	31.048	13.353
1072	C	TYR	127	41.296	31.731	14.254
1073	O	TYR	127	41.480	32.890	14.643
1074	CB	TYR	127	43.312	30.279	14.223
1075	CG	TYR	127	43.837	30.968	15.484
1076	CD1	TYR	127	43.322	30.600	16.722
1077	CD2	TYR	127	44.840	31.928	15.409
1078	CE1	TYR	127	43.794	31.203	17.880
1079	CE2	TYR	127	45.314	32.531	16.566
1080	CZ	TYR	127	44.789	32.168	17.799
1081	OH	TYR	127	45.274	32.748	18.950
1082	N	GLY	128	40.221	31.017	14.535
1083	CA	GLY	128	39.213	31.485	15.492
1084	C	GLY	128	38.851	30.389	16.496
1085	O	GLY	128	39.496	29.332	16.556
1086	N	ASN	129	37.833	30.661	17.297
1087	CA	ASN	129	37.370	29.698	18.319
1088	C	ASN	129	35.845	29.721	18.439
1089	O	ASN	129	35.277	30.678	18.980
1090	CB	ASN	129	37.929	30.056	19.699
1091	CG	ASN	129	39.449	30.202	19.728

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1092	OD1	ASN	129	39.965	31.321	19.633
1093	ND2	ASN	129	40.135	29.094	19.951
1094	N	GLY	130	35.199	28.642	18.039
1095	CA	GLY	130	33.734	28.584	18.122
1096	C	GLY	130	33.259	27.661	19.242
1097	O	GLY	130	33.769	26.546	19.409
1098	N	THR	131	32.330	28.153	20.044
1099	CA	THR	131	31.795	27.327	21.137
1100	C	THR	131	30.557	26.540	20.702
1101	O	THR	131	29.553	27.095	20.229
1102	CB	THR	131	31.481	28.208	22.336
1103	OG1	THR	131	30.437	29.095	21.981
1104	CG2	THR	131	32.693	29.035	22.747
1105	N	ARG	132	30.660	25.237	20.897
1106	CA	ARG	132	29.653	24.267	20.451
1107	C	ARG	132	28.816	23.736	21.613
1108	O	ARG	132	29.295	22.920	22.410
1109	CB	ARG	132	30.435	23.093	19.865
1110	CG	ARG	132	31.490	23.561	18.867
1111	CD	ARG	132	30.849	23.917	17.537
1112	NE	ARG	132	31.484	25.081	16.914
1113	CZ	ARG	132	30.748	26.119	16.528
1114	NH1	ARG	132	29.431	26.097	16.733
1115	NH2	ARG	132	31.323	27.173	15.947
1116	N	LYS	133	27.551	24.114	21.653
1117	CA	LYS	133	26.682	23.638	22.731
1118	C	LYS	133	25.874	22.419	22.290
1119	O	LYS	133	24.861	22.535	21.588
1120	CB	LYS	133	25.758	24.771	23.139
1121	CG	LYS	133	25.132	24.479	24.492
1122	CD	LYS	133	24.312	25.663	24.977
1123	CE	LYS	133	23.949	25.500	26.445
1124	NZ	LYS	133	25.168	25.439	27.269
1125	N	SER	134	26.329	21.266	22.751
1126	CA	SER	134	25.747	19.968	22.380
1127	C	SER	134	24.587	19.579	23.312
1128	O	SER	134	24.332	20.285	24.296
1129	CB	SER	134	26.889	18.957	22.461
1130	OG	SER	134	27.937	19.407	21.614

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1131	N	PRO	135	23.850	18.523	22.981
1132	CA	PRO	135	22.818	18.009	23.894
1133	C	PRO	135	23.446	17.471	25.173
1134	O	PRO	135	24.550	16.917	25.147
1135	CB	PRO	135	22.114	16.922	23.152
1136	CG	PRO	135	22.863	16.649	21.863
1137	CD	PRO	135	23.972	17.684	21.783
1138	N	GLU	136	22.727	17.670	26.273
1139	CA	GLU	136	23.207	17.398	27.644
1140	C	GLU	136	24.326	18.394	27.965
1141	O	GLU	136	25.226	18.134	28.774
1142	CB	GLU	136	23.675	15.943	27.750
1143	CG	GLU	136	23.964	15.490	29.178
1144	CD	GLU	136	24.465	14.050	29.167
1145	OE1	GLU	136	23.967	13.287	28.352
1146	OE2	GLU	136	25.248	13.711	30.045
1147	N	MET	137	24.117	19.587	27.427
1148	CA	MET	137	25.066	20.716	27.349
1149	C	MET	137	26.547	20.494	27.697
1150	O	MET	137	27.026	21.140	28.637
1151	CB	MET	137	24.496	21.788	28.265
1152	CG	MET	137	23.049	22.083	27.883
1153	SD	MET	137	22.171	23.229	28.969
1154	CE	MET	137	22.263	22.281	30.505
1155	N	PRO	138	27.289	19.667	26.966
1156	CA	PRO	138	28.733	19.844	26.939
1157	C	PRO	138	29.093	21.001	26.013
1158	O	PRO	138	28.664	21.051	24.851
1159	CB	PRO	138	29.261	18.553	26.397
1160	CG	PRO	138	28.118	17.834	25.699
1161	CD	PRO	138	26.894	18.705	25.937
1162	N	VAL	139	29.824	21.954	26.561
1163	CA	VAL	139	30.360	23.043	25.742
1164	C	VAL	139	31.705	22.622	25.159
1165	O	VAL	139	32.728	22.590	25.855
1166	CB	VAL	139	30.517	24.291	26.603
1167	CG1	VAL	139	31.062	25.457	25.782
1168	CG2	VAL	139	29.185	24.671	27.242
1169	N	ILE	140	31.659	22.214	23.905

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1170	CA	ILE	140	32.854	21.758	23.193
1171	C	ILE	140	33.541	22.963	22.566
1172	O	ILE	140	32.875	23.819	21.974
1173	CB	ILE	140	32.402	20.782	22.106
1174	CG1	ILE	140	31.567	19.653	22.702
1175	CG2	ILE	140	33.588	20.200	21.341
1176	CD1	ILE	140	32.388	18.793	23.657
1177	N	LEU	141	34.839	23.096	22.765
1178	CA	LEU	141	35.519	24.222	22.121
1179	C	LEU	141	36.142	23.789	20.799
1180	O	LEU	141	36.929	22.835	20.726
1181	CB	LEU	141	36.556	24.824	23.058
1182	CG	LEU	141	35.905	25.372	24.325
1183	CD1	LEU	141	36.950	25.865	25.320
1184	CD2	LEU	141	34.918	26.486	23.997
1185	N	GLU	142	35.701	24.447	19.744
1186	CA	GLU	142	36.220	24.162	18.410
1187	C	GLU	142	37.230	25.216	17.980
1188	O	GLU	142	36.867	26.322	17.557
1189	CB	GLU	142	35.063	24.142	17.421
1190	CG	GLU	142	35.535	23.829	16.006
1191	CD	GLU	142	34.364	23.969	15.042
1192	OE1	GLU	142	34.383	24.899	14.248
1193	OE2	GLU	142	33.430	23.193	15.183
1194	N	VAL	143	38.497	24.865	18.103
1195	CA	VAL	143	39.551	25.740	17.593
1196	C	VAL	143	39.539	25.597	16.077
1197	O	VAL	143	39.555	24.479	15.548
1198	CB	VAL	143	40.886	25.321	18.195
1199	CG1	VAL	143	42.008	26.264	17.774
1200	CG2	VAL	143	40.785	25.277	19.716
1201	N	SER	144	39.387	26.715	15.394
1202	CA	SER	144	39.135	26.669	13.957
1203	C	SER	144	40.153	27.472	13.162
1204	O	SER	144	39.933	28.647	12.838
1205	CB	SER	144	37.738	27.228	13.723
1206	OG	SER	144	36.846	26.492	14.551
1207	N	LEU	145	41.249	26.817	12.833
1208	CA	LEU	145	42.274	27.425	11.988

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1209	C	LEU	145	41.905	27.254	10.528
1210	O	LEU	145	41.200	26.309	10.154
1211	CB	LEU	145	43.638	26.767	12.183
1212	CG	LEU	145	44.354	27.089	13.488
1213	CD1	LEU	145	43.968	26.140	14.621
1214	CD2	LEU	145	45.854	26.999	13.250
1215	N	GLU	146	42.300	28.227	9.733
1216	CA	GLU	146	42.161	28.128	8.282
1217	C	GLU	146	43.514	28.333	7.607
1218	O	GLU	146	43.666	28.098	6.402
1219	CB	GLU	146	41.160	29.173	7.809
1220	CG	GLU	146	39.783	28.912	8.407
1221	CD	GLU	146	38.818	30.004	7.973
1222	OE1	GLU	146	37.634	29.860	8.243
1223	OE2	GLU	146	39.302	31.022	7.498
1224	N	GLY	147	44.509	28.695	8.403
1225	CA	GLY	147	45.857	28.936	7.865
1226	C	GLY	147	46.783	27.717	7.904
1227	O	GLY	147	47.846	27.756	8.542
1228	N	SER	148	46.399	26.677	7.182
1229	CA	SER	148	47.225	25.468	7.052
1230	C	SER	148	47.037	24.825	5.679
1231	O	SER	148	45.961	24.298	5.371
1232	CB	SER	148	46.828	24.468	8.127
1233	OG	SER	148	47.557	23.273	7.884
1234	N	HIS	149	48.097	24.816	4.889
1235	CA	HIS	149	47.988	24.275	3.530
1236	C	HIS	149	48.230	22.767	3.514
1237	O	HIS	149	49.224	22.265	4.054
1238	CB	HIS	149	49.003	24.975	2.632
1239	CG	HIS	149	48.706	24.855	1.146
1240	ND1	HIS	149	49.112	23.883	0.306
1241	CD2	HIS	149	47.952	25.735	0.407
1242	CE1	HIS	149	48.640	24.141	-0.932
1243	NE2	HIS	149	47.922	25.284	-0.868
1244	N	ASP	150	47.331	22.062	2.844
1245	CA	ASP	150	47.478	20.610	2.653
1246	C	ASP	150	48.779	20.266	1.937
1247	O	ASP	150	49.322	21.082	1.181

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1248	CB	ASP	150	46.288	20.059	1.875
1249	CG	ASP	150	45.953	20.909	0.651
1250	OD1	ASP	150	44.799	21.314	0.573
1251	OD2	ASP	150	46.803	21.055	-0.215
1252	N	THR	151	49.334	19.128	2.332
1253	CA	THR	151	50.624	18.579	1.860
1254	C	THR	151	51.790	19.545	2.063
1255	O	THR	151	52.779	19.498	1.322
1256	CB	THR	151	50.543	18.139	0.395
1257	OG1	THR	151	50.502	19.266	-0.471
1258	CG2	THR	151	49.326	17.259	0.120
1259	N	ALA	152	51.714	20.328	3.127
1260	CA	ALA	152	52.736	21.327	3.414
1261	C	ALA	152	52.840	21.573	4.912
1262	O	ALA	152	52.810	20.644	5.728
1263	CB	ALA	152	52.372	22.621	2.692
1264	N	ASN	153	52.985	22.839	5.259
1265	CA	ASN	153	53.190	23.224	6.656
1266	C	ASN	153	52.165	24.255	7.117
1267	O	ASN	153	51.289	24.707	6.365
1268	CB	ASN	153	54.603	23.786	6.811
1269	CG	ASN	153	55.656	22.760	6.391
1270	OD1	ASN	153	55.979	22.630	5.204
1271	ND2	ASN	153	56.181	22.045	7.369
1272	N	CYS	154	52.238	24.525	8.410
1273	CA	CYS	154	51.447	25.591	9.052
1274	C	CYS	154	52.138	26.939	8.778
1275	O	CYS	154	53.232	26.966	8.198
1276	CB	CYS	154	51.391	25.286	10.545
1277	SG	CYS	154	50.054	26.034	11.509
1278	N	GLU	155	51.557	28.025	9.264
1279	CA	GLU	155	52.117	29.369	9.054
1280	C	GLU	155	53.257	29.786	9.990
1281	O	GLU	155	53.626	30.965	9.965
1282	CB	GLU	155	51.005	30.393	9.203
1283	CG	GLU	155	49.902	30.219	8.173
1284	CD	GLU	155	48.859	31.300	8.412
1285	OE1	GLU	155	48.885	31.854	9.504
1286	OE2	GLU	155	48.054	31.545	7.523

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1287	N	ALA	156	53.771	28.842	10.774
1288	CA	ALA	156	54.873	29.002	11.753
1289	C	ALA	156	55.556	30.365	11.780
1290	O	ALA	156	55.116	31.246	12.528
1291	CB	ALA	156	55.907	27.932	11.432
1292	N	CYS	157	56.718	30.467	11.168
1293	CA	CYS	157	57.322	31.789	10.973
1294	C	CYS	157	57.628	31.967	9.495
1295	O	CYS	157	58.232	31.082	8.879
1296	CB	CYS	157	58.582	31.940	11.807
1297	SG	CYS	157	59.396	33.553	11.789
1298	N	THR	158	57.244	33.125	8.975
1299	CA	THR	158	57.314	33.456	7.536
1300	C	THR	158	56.842	32.270	6.694
1301	O	THR	158	57.615	31.490	6.120
1302	CB	THR	158	58.701	33.956	7.152
1303	OG1	THR	158	59.671	32.958	7.448
1304	CG2	THR	158	59.058	35.212	7.954
1305	N	LEU	159	55.528	32.166	6.647
1306	CA	LEU	159	54.855	31.022	6.027
1307	C	LEU	159	55.230	30.810	4.567
1308	O	LEU	159	55.449	31.761	3.807
1309	CB	LEU	159	53.352	31.254	6.171
1310	CG	LEU	159	52.901	32.657	5.726
1311	CD1	LEU	159	52.489	32.685	4.255
1312	CD2	LEU	159	51.731	33.142	6.574
1313	N	GLY	160	55.479	29.554	4.245
1314	CA	GLY	160	55.725	29.173	2.854
1315	C	GLY	160	57.194	28.919	2.520
1316	O	GLY	160	57.495	28.403	1.434
1317	N	ILE	161	58.096	29.259	3.425
1318	CA	ILE	161	59.517	29.046	3.138
1319	C	ILE	161	59.962	27.637	3.532
1320	O	ILE	161	60.311	27.375	4.689
1321	CB	ILE	161	60.322	30.107	3.877
1322	CG1	ILE	161	59.866	31.496	3.449
1323	CG2	ILE	161	61.815	29.938	3.611
1324	CD1	ILE	161	60.703	32.590	4.102
1325	N	CYS	162	59.837	26.720	2.585

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1326	CA	CYS	162	60.240	25.323	2.805
1327	C	CYS	162	60.967	24.755	1.589
1328	O	CYS	162	62.167	24.978	1.399
1329	CB	CYS	162	59.002	24.479	3.096
1330	SG	CYS	162	58.093	24.908	4.599
1331	N	GLY	163	60.241	23.965	0.816
1332	CA	GLY	163	60.820	23.335	-0.375
1333	C	GLY	163	59.815	22.410	-1.052
1334	O	GLY	163	59.597	21.279	-0.604
1335	N	GLN	164	59.385	22.824	-2.234
1336	CA	GLN	164	58.346	22.120	-3.010
1337	C	GLN	164	58.861	20.968	-3.885
1338	O	GLN	164	58.222	20.642	-4.894
1339	CB	GLN	164	57.686	23.131	-3.937
1340	CG	GLN	164	58.725	23.759	-4.858
1341	CD	GLN	164	58.086	24.084	-6.201
1342	OE1	GLN	164	57.945	25.253	-6.577
1343	NE2	GLN	164	57.706	23.030	-6.903
1344	N	GLY	165	60.022	20.423	-3.568
1345	CA	GLY	165	60.610	19.387	-4.418
1346	C	GLY	165	60.494	18.015	-3.769
1347	O	GLY	165	59.534	17.273	-4.009
1348	N	LEU	166	61.509	17.669	-3.000
1349	CA	LEU	166	61.542	16.361	-2.348
1350	C	LEU	166	62.090	16.513	-0.934
1351	O	LEU	166	61.390	16.314	0.065
1352	CB	LEU	166	62.455	15.459	-3.175
1353	CG	LEU	166	62.340	13.992	-2.779
1354	CD1	LEU	166	60.932	13.468	-3.044
1355	CD2	LEU	166	63.372	13.148	-3.517
1356	N	LYS	167	63.362	16.861	-0.869
1357	CA	LYS	167	64.007	17.110	0.417
1358	C	LYS	167	63.762	18.551	0.834
1359	O	LYS	167	63.826	19.470	0.009
1360	CB	LYS	167	65.498	16.860	0.251
1361	CG	LYS	167	65.763	15.469	-0.306
1362	CD	LYS	167	67.234	15.283	-0.649
1363	CE	LYS	167	67.503	13.886	-1.190
1364	NZ	LYS	167	68.925	13.725	-1.531

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1365	N	SER	168	63.423	18.738	2.095
1366	CA	SER	168	63.231	20.095	2.607
1367	C	SER	168	64.118	20.306	3.830
1368	O	SER	168	65.155	19.645	3.960
1369	CB	SER	168	61.760	20.285	2.954
1370	OG	SER	168	61.371	21.586	2.533
1371	N	CYS	169	63.788	21.313	4.622
1372	CA	CYS	169	64.514	21.597	5.860
1373	C	CYS	169	63.550	21.796	7.028
1374	O	CYS	169	62.526	21.113	7.146
1375	CB	CYS	169	65.357	22.856	5.673
1376	SG	CYS	169	66.669	22.759	4.432
1377	N	MET	170	63.951	22.686	7.921
1378	CA	MET	170	63.156	23.070	9.100
1379	C	MET	170	62.296	24.303	8.802
1380	O	MET	170	62.108	24.597	7.618
1381	CB	MET	170	64.148	23.388	10.213
1382	CG	MET	170	65.205	24.370	9.710
1383	SD	MET	170	66.511	24.808	10.878
1384	CE	MET	170	67.479	25.898	9.808
1385	N	THR	171	61.611	24.797	9.838
1386	CA	THR	171	60.888	26.111	9.909
1387	C	THR	171	59.694	26.062	10.881
1388	O	THR	171	58.635	25.499	10.583
1389	CB	THR	171	60.427	26.691	8.561
1390	OG1	THR	171	61.565	27.244	7.908
1391	CG2	THR	171	59.456	27.856	8.741
1392	N	LYS	172	59.921	26.619	12.063
1393	CA	LYS	172	58.897	26.794	13.118
1394	C	LYS	172	58.899	28.290	13.489
1395	O	LYS	172	58.997	29.075	12.544
1396	CB	LYS	172	59.305	25.893	14.281
1397	CG	LYS	172	59.120	24.418	13.974
1398	CD	LYS	172	59.406	23.599	15.226
1399	CE	LYS	172	59.152	22.117	14.994
1400	NZ	LYS	172	59.351	21.356	16.238
1401	N	PRO	173	58.735	28.713	14.742
1402	CA	PRO	173	57.766	28.214	15.758
1403	C	PRO	173	56.388	28.909	15.962

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1404	O	PRO	173	55.414	28.216	16.296
1405	CB	PRO	173	58.560	28.468	17.006
1406	CG	PRO	173	59.496	29.652	16.740
1407	CD	PRO	173	59.399	29.918	15.244
1408	N	SER	174	56.254	30.182	15.618
1409	CA	SER	174	55.260	31.059	16.269
1410	C	SER	174	53.809	30.677	16.039
1411	O	SER	174	53.195	30.074	16.930
1412	CB	SER	174	55.488	32.490	15.807
1413	OG	SER	174	56.764	32.900	16.283
1414	N	LYS	175	53.376	30.769	14.795
1415	CA	LYS	175	51.970	30.560	14.437
1416	C	LYS	175	51.518	29.096	14.425
1417	O	LYS	175	50.373	28.818	14.055
1418	CB	LYS	175	51.754	31.157	13.058
1419	CG	LYS	175	52.184	32.617	13.029
1420	CD	LYS	175	52.195	33.151	11.604
1421	CE	LYS	175	52.753	34.565	11.541
1422	NZ	LYS	175	52.838	35.023	10.146
1423	N	SER	176	52.391	28.172	14.788
1424	CA	SER	176	51.959	26.792	14.942
1425	C	SER	176	51.703	26.495	16.413
1426	O	SER	176	50.684	25.887	16.751
1427	CB	SER	176	53.044	25.862	14.420
1428	OG	SER	176	52.583	24.538	14.635
1429	N	LEU	177	52.502	27.103	17.276
1430	CA	LEU	177	52.394	26.823	18.710
1431	C	LEU	177	51.396	27.738	19.408
1432	O	LEU	177	50.724	27.297	20.350
1433	CB	LEU	177	53.771	26.984	19.341
1434	CG	LEU	177	54.765	25.992	18.747
1435	CD1	LEU	177	56.171	26.245	19.280
1436	CD2	LEU	177	54.334	24.553	19.012
1437	N	LEU	178	51.149	28.902	18.832
1438	CA	LEU	178	50.118	29.811	19.370
1439	C	LEU	178	48.721	29.171	19.506
1440	O	LEU	178	48.258	29.088	20.653
1441	CB	LEU	178	50.050	31.084	18.531
1442	CG	LEU	178	51.312	31.923	18.688

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1443	CD1	LEU	178	51.300	33.119	17.743
1444	CD2	LEU	178	51.496	32.375	20.133
1445	N	PRO	179	48.122	28.579	18.471
1446	CA	PRO	179	46.816	27.930	18.672
1447	C	PRO	179	46.848	26.611	19.465
1448	O	PRO	179	45.781	26.118	19.841
1449	CB	PRO	179	46.285	27.681	17.296
1450	CG	PRO	179	47.380	27.943	16.278
1451	CD	PRO	179	48.560	28.466	17.067
1452	N	HIS	180	48.021	26.095	19.804
1453	CA	HIS	180	48.089	24.881	20.622
1454	C	HIS	180	48.195	25.262	22.094
1455	O	HIS	180	47.873	24.460	22.978
1456	CB	HIS	180	49.306	24.047	20.229
1457	CG	HIS	180	49.308	23.530	18.803
1458	ND1	HIS	180	48.236	23.324	18.012
1459	CD2	HIS	180	50.416	23.165	18.076
1460	CE1	HIS	180	48.648	22.871	16.812
1461	NE2	HIS	180	49.995	22.770	16.853
1462	N	LEU	181	48.567	26.508	22.339
1463	CA	LEU	181	48.602	27.033	23.702
1464	C	LEU	181	47.289	27.722	24.058
1465	O	LEU	181	46.948	27.823	25.242
1466	CB	LEU	181	49.755	28.023	23.808
1467	CG	LEU	181	51.095	27.339	23.561
1468	CD1	LEU	181	52.224	28.357	23.449
1469	CD2	LEU	181	51.396	26.305	24.642
1470	N	LYS	182	46.491	28.038	23.049
1471	CA	LYS	182	45.169	28.639	23.292
1472	C	LYS	182	44.065	27.571	23.300
1473	O	LYS	182	43.075	27.648	22.560
1474	CB	LYS	182	44.894	29.696	22.227
1475	CG	LYS	182	43.714	30.576	22.627
1476	CD	LYS	182	43.409	31.638	21.581
1477	CE	LYS	182	42.225	32.497	22.008
1478	NZ	LYS	182	41.952	33.546	21.014
1479	N	THR	183	44.269	26.567	24.137
1480	CA	THR	183	43.336	25.442	24.273
1481	C	THR	183	43.214	25.022	25.736

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1482	O	THR	183	43.365	25.829	26.659
1483	CB	THR	183	43.846	24.250	23.463
1484	OG1	THR	183	45.214	24.061	23.792
1485	CG2	THR	183	43.756	24.465	21.956
1486	N	GLY	184	42.857	23.762	25.919
1487	CA	GLY	184	42.775	23.154	27.256
1488	C	GLY	184	41.931	21.889	27.176
1489	O	GLY	184	41.832	21.103	28.128
1490	N	ASN	185	41.618	21.569	25.933
1491	CA	ASN	185	40.648	20.531	25.561
1492	C	ASN	185	41.230	19.117	25.548
1493	O	ASN	185	40.532	18.141	25.240
1494	CB	ASN	185	40.195	20.873	24.149
1495	CG	ASN	185	39.920	22.369	24.019
1496	OD1	ASN	185	40.771	23.116	23.518
1497	ND2	ASN	185	38.807	22.808	24.577
1498	N	SER	186	42.503	19.034	25.894
1499	CA	SER	186	43.246	17.778	25.885
1500	C	SER	186	43.497	17.260	27.297
1501	O	SER	186	44.223	16.274	27.462
1502	CB	SER	186	44.583	18.026	25.200
1503	OG	SER	186	45.313	18.948	25.999
1504	N	SER	187	42.968	17.954	28.295
1505	CA	SER	187	43.157	17.537	29.692
1506	C	SER	187	42.603	16.134	29.926
1507	O	SER	187	41.422	15.871	29.662
1508	CB	SER	187	42.442	18.534	30.599
1509	OG	SER	187	43.060	19.802	30.424
1510	N	PRO	188	43.451	15.276	30.480
1511	CA	PRO	188	43.231	13.816	30.490
1512	C	PRO	188	42.244	13.328	31.557
1513	O	PRO	188	42.624	12.661	32.526
1514	CB	PRO	188	44.585	13.225	30.737
1515	CG	PRO	188	45.556	14.332	31.113
1516	CD	PRO	188	44.789	15.634	30.959
1517	N	GLY	189	40.983	13.649	31.339
1518	CA	GLY	189	39.884	13.226	32.204
1519	C	GLY	189	38.612	13.275	31.370
1520	O	GLY	189	37.663	12.511	31.583

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1521	N	ILE	190	38.620	14.196	30.424
1522	CA	ILE	190	37.528	14.324	29.459
1523	C	ILE	190	37.892	13.468	28.238
1524	O	ILE	190	39.053	13.044	28.135
1525	CB	ILE	190	37.408	15.812	29.120
1526	CG1	ILE	190	36.073	16.164	28.464
1527	CG2	ILE	190	38.574	16.257	28.243
1528	CD1	ILE	190	34.895	15.836	29.374
1529	N	GLY	191	36.916	13.141	27.400
1530	CA	GLY	191	37.166	12.348	26.186
1531	C	GLY	191	38.315	12.904	25.350
1532	O	GLY	191	38.544	14.122	25.295
1533	N	ALA	192	39.028	11.984	24.717
1534	CA	ALA	192	40.223	12.288	23.916
1535	C	ALA	192	39.990	13.412	22.917
1536	O	ALA	192	39.026	13.381	22.145
1537	CB	ALA	192	40.620	11.025	23.160
1538	N	VAL	193	40.906	14.365	22.909
1539	CA	VAL	193	40.760	15.563	22.071
1540	C	VAL	193	40.772	15.214	20.581
1541	O	VAL	193	41.678	14.536	20.071
1542	CB	VAL	193	41.884	16.528	22.442
1543	CG1	VAL	193	43.224	15.808	22.516
1544	CG2	VAL	193	41.947	17.749	21.527
1545	N	TYR	194	39.709	15.630	19.911
1546	CA	TYR	194	39.529	15.303	18.499
1547	C	TYR	194	40.231	16.287	17.581
1548	O	TYR	194	40.231	17.505	17.811
1549	CB	TYR	194	38.042	15.283	18.187
1550	CG	TYR	194	37.528	13.892	17.847
1551	CD1	TYR	194	37.321	12.964	18.859
1552	CD2	TYR	194	37.287	13.547	16.523
1553	CE1	TYR	194	36.863	11.691	18.549
1554	CE2	TYR	194	36.826	12.275	16.212
1555	CZ	TYR	194	36.615	11.351	17.227
1556	OH	TYR	194	36.167	10.085	16.919
1557	N	LEU	195	40.900	15.719	16.591
1558	CA	LEU	195	41.550	16.507	15.541
1559	C	LEU	195	41.062	16.033	14.175

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1560	O	LEU	195	41.683	15.165	13.541
1561	CB	LEU	195	43.064	16.322	15.614
1562	CG	LEU	195	43.662	16.706	16.965
1563	CD1	LEU	195	45.169	16.544	16.930
1564	CD2	LEU	195	43.327	18.137	17.359
1565	N	ALA	196	39.918	16.554	13.763
1566	CA	ALA	196	39.370	16.213	12.445
1567	C	ALA	196	40.278	16.750	11.344
1568	O	ALA	196	41.037	17.702	11.553
1569	CB	ALA	196	37.969	16.785	12.308
1570	N	ASN	197	40.339	16.019	10.249
1571	CA	ASN	197	41.320	16.347	9.212
1572	C	ASN	197	40.750	16.342	7.810
1573	O	ASN	197	39.701	15.755	7.537
1574	CB	ASN	197	42.424	15.303	9.272
1575	CG	ASN	197	43.755	15.954	9.615
1576	OD1	ASN	197	44.756	15.798	8.895
1577	ND2	ASN	197	43.699	16.785	10.638
1578	N	GLN	198	41.515	16.929	6.900
1579	CA	GLN	198	41.183	16.844	5.476
1580	C	GLN	198	41.548	15.456	4.992
1581	O	GLN	198	40.824	14.867	4.190
1582	CB	GLN	198	41.969	17.875	4.674
1583	CG	GLN	198	41.331	19.260	4.694
1584	CD	GLN	198	42.254	20.264	4.013
1585	OE1	GLN	198	42.519	21.346	4.548
1586	NE2	GLN	198	42.760	19.878	2.856
1587	N	ALA	199	42.612	14.912	5.558
1588	CA	ALA	199	42.944	13.506	5.353
1589	C	ALA	199	42.734	12.795	6.678
1590	O	ALA	199	41.610	12.444	7.046
1591	CB	ALA	199	44.400	13.381	4.916
1592	N	LYS	200	43.830	12.640	7.396
1593	CA	LYS	200	43.810	12.052	8.741
1594	C	LYS	200	45.172	12.247	9.389
1595	O	LYS	200	45.316	12.585	10.570
1596	CB	LYS	200	43.539	10.562	8.596
1597	CG	LYS	200	43.701	9.819	9.921
1598	CD	LYS	200	43.726	8.305	9.746

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1599	CE	LYS	200	45.069	7.732	9.282
1600	NZ	LYS	200	45.372	7.956	7.859
1601	N	ASN	201	46.164	12.100	8.535
1602	CA	ASN	201	47.568	12.120	8.922
1603	C	ASN	201	48.146	13.509	9.161
1604	O	ASN	201	48.036	14.033	10.275
1605	CB	ASN	201	48.363	11.410	7.826
1606	CG	ASN	201	47.686	11.435	6.455
1607	OD1	ASN	201	47.408	12.500	5.882
1608	ND2	ASN	201	47.493	10.247	5.920
1609	N	GLN	202	48.569	14.138	8.078
1610	CA	GLN	202	49.451	15.317	8.074
1611	C	GLN	202	49.271	16.292	9.221
1612	O	GLN	202	50.008	16.217	10.214
1613	CB	GLN	202	49.217	16.045	6.766
1614	CG	GLN	202	49.431	15.087	5.608
1615	CD	GLN	202	49.357	15.850	4.301
1616	OE1	GLN	202	48.923	17.009	4.266
1617	NE2	GLN	202	49.847	15.209	3.255
1618	N	SER	203	48.197	17.058	9.182
1619	CA	SER	203	48.049	18.130	10.163
1620	C	SER	203	47.592	17.649	11.540
1621	O	SER	203	48.048	18.225	12.534
1622	CB	SER	203	47.080	19.168	9.614
1623	OG	SER	203	45.840	18.530	9.362
1624	N	ALA	204	46.974	16.480	11.623
1625	CA	ALA	204	46.511	16.004	12.929
1626	C	ALA	204	47.684	15.449	13.701
1627	O	ALA	204	47.906	15.854	14.847
1628	CB	ALA	204	45.493	14.887	12.757
1629	N	GLU	205	48.580	14.811	12.970
1630	CA	GLU	205	49.760	14.209	13.576
1631	C	GLU	205	50.875	15.224	13.773
1632	O	GLU	205	51.689	15.049	14.683
1633	CB	GLU	205	50.213	13.090	12.663
1634	CG	GLU	205	49.071	12.108	12.433
1635	CD	GLU	205	48.807	11.253	13.666
1636	OE1	GLU	205	49.781	10.912	14.323
1637	OE2	GLU	205	47.687	10.774	13.785

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1638	N	ALA	206	50.775	16.364	13.111
1639	CA	ALA	206	51.700	17.457	13.407
1640	C	ALA	206	51.337	18.075	14.752
1641	O	ALA	206	52.191	18.122	15.651
1642	CB	ALA	206	51.593	18.510	12.310
1643	N	LYS	207	50.040	18.226	14.978
1644	CA	LYS	207	49.556	18.769	16.251
1645	C	LYS	207	49.750	17.769	17.389
1646	O	LYS	207	50.224	18.149	18.467
1647	CB	LYS	207	48.066	19.086	16.123
1648	CG	LYS	207	47.776	20.086	15.007
1649	CD	LYS	207	46.357	20.644	15.107
1650	CE	LYS	207	45.270	19.634	14.750
1651	NZ	LYS	207	45.137	19.454	13.296
1652	N	GLU	208	49.648	16.491	17.065
1653	CA	GLU	208	49.827	15.435	18.062
1654	C	GLU	208	51.289	15.160	18.399
1655	O	GLU	208	51.582	14.881	19.568
1656	CB	GLU	208	49.171	14.174	17.515
1657	CG	GLU	208	47.662	14.361	17.495
1658	CD	GLU	208	46.963	13.323	16.627
1659	OE1	GLU	208	45.847	12.949	16.963
1660	OE2	GLU	208	47.505	13.018	15.574
1661	N	ALA	209	52.197	15.438	17.478
1662	CA	ALA	209	53.616	15.215	17.750
1663	C	ALA	209	54.222	16.370	18.530
1664	O	ALA	209	55.003	16.124	19.456
1665	CB	ALA	209	54.363	15.043	16.432
1666	N	LYS	210	53.723	17.577	18.321
1667	CA	LYS	210	54.224	18.690	19.131
1668	C	LYS	210	53.534	18.736	20.490
1669	O	LYS	210	54.209	18.952	21.507
1670	CB	LYS	210	54.044	20.002	18.379
1671	CG	LYS	210	55.010	20.065	17.198
1672	CD	LYS	210	55.019	21.424	16.500
1673	CE	LYS	210	54.219	21.438	15.199
1674	NZ	LYS	210	52.777	21.257	15.422
1675	N	GLY	211	52.293	18.279	20.527
1676	CA	GLY	211	51.554	18.163	21.785

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1677	C	GLY	211	52.206	17.135	22.699
1678	O	GLY	211	52.699	17.491	23.777
1679	N	SER	212	52.407	15.938	22.172
1680	CA	SER	212	53.013	14.857	22.953
1681	C	SER	212	54.494	15.090	23.237
1682	O	SER	212	54.953	14.789	24.346
1683	CB	SER	212	52.856	13.571	22.158
1684	OG	SER	212	51.468	13.276	22.098
1685	N	GLY	213	55.168	15.794	22.342
1686	CA	GLY	213	56.561	16.185	22.555
1687	C	GLY	213	56.701	17.075	23.782
1688	O	GLY	213	57.360	16.682	24.754
1689	N	TYR	214	55.911	18.136	23.836
1690	CA	TYR	214	56.001	19.069	24.965
1691	C	TYR	214	55.356	18.535	26.244
1692	O	TYR	214	55.811	18.890	27.337
1693	CB	TYR	214	55.337	20.385	24.575
1694	CG	TYR	214	56.053	21.139	23.458
1695	CD1	TYR	214	57.442	21.175	23.425
1696	CD2	TYR	214	55.317	21.802	22.484
1697	CE1	TYR	214	58.097	21.859	22.409
1698	CE2	TYR	214	55.970	22.486	21.468
1699	CZ	TYR	214	57.357	22.512	21.433
1700	OH	TYR	214	58.002	23.209	20.436
1701	N	GLU	215	54.489	17.541	26.126
1702	CA	GLU	215	53.903	16.926	27.321
1703	C	GLU	215	54.825	15.867	27.928
1704	O	GLU	215	54.740	15.591	29.128
1705	CB	GLU	215	52.582	16.276	26.937
1706	CG	GLU	215	51.858	15.731	28.162
1707	CD	GLU	215	50.651	14.926	27.711
1708	OE1	GLU	215	50.707	14.430	26.590
1709	OE2	GLU	215	49.637	14.984	28.389
1710	N	LYS	216	55.758	15.364	27.139
1711	CA	LYS	216	56.773	14.447	27.654
1712	C	LYS	216	57.967	15.237	28.178
1713	O	LYS	216	58.524	14.924	29.240
1714	CB	LYS	216	57.206	13.579	26.478
1715	CG	LYS	216	58.363	12.646	26.809

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1716	CD	LYS	216	58.842	11.951	25.540
1717	CE	LYS	216	60.046	11.052	25.788
1718	NZ	LYS	216	60.487	10.436	24.526
1719	N	LEU	217	58.175	16.389	27.563
1720	CA	LEU	217	59.301	17.248	27.917
1721	C	LEU	217	59.034	18.013	29.206
1722	O	LEU	217	59.947	18.167	30.028
1723	CB	LEU	217	59.492	18.240	26.782
1724	CG	LEU	217	60.964	18.561	26.610
1725	CD1	LEU	217	61.709	17.270	26.305
1726	CD2	LEU	217	61.172	19.581	25.497
1727	N	GLY	218	57.769	18.342	29.423
1728	CA	GLY	218	57.291	18.984	30.659
1729	C	GLY	218	57.804	18.304	31.924
1730	O	GLY	218	58.811	18.753	32.483
1731	N	PRO	219	57.189	17.198	32.321
1732	CA	PRO	219	57.581	16.507	33.559
1733	C	PRO	219	58.941	15.794	33.524
1734	O	PRO	219	59.372	15.299	34.570
1735	CB	PRO	219	56.494	15.509	33.812
1736	CG	PRO	219	55.544	15.485	32.627
1737	CD	PRO	219	56.027	16.573	31.685
1738	N	SER	220	59.615	15.745	32.384
1739	CA	SER	220	60.965	15.180	32.368
1740	C	SER	220	62.009	16.264	32.615
1741	O	SER	220	63.120	15.950	33.062
1742	CB	SER	220	61.240	14.531	31.013
1743	OG	SER	220	61.303	15.550	30.021
1744	N	ARG	221	61.603	17.516	32.423
1745	CA	ARG	221	62.469	18.704	32.519
1746	C	ARG	221	63.838	18.451	31.900
1747	O	ARG	221	64.877	18.632	32.546
1748	CB	ARG	221	62.603	19.102	33.983
1749	CG	ARG	221	61.226	19.289	34.609
1750	CD	ARG	221	61.323	19.897	36.002
1751	NE	ARG	221	61.916	21.242	35.933
1752	CZ	ARG	221	61.196	22.366	35.985
1753	NH1	ARG	221	61.801	23.550	35.864
1754	NH2	ARG	221	59.869	22.307	36.119

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1755	N	ASP	222	63.820	18.059	30.636
1756	CA	ASP	222	65.043	17.604	29.963
1757	C	ASP	222	64.751	17.349	28.488
1758	O	ASP	222	64.048	16.396	28.138
1759	CB	ASP	222	65.501	16.314	30.654
1760	CG	ASP	222	66.843	15.792	30.135
1761	OD1	ASP	222	67.334	16.357	29.165
1762	OD2	ASP	222	67.184	14.685	30.521
1763	N	PRO	223	65.326	18.184	27.636
1764	CA	PRO	223	65.065	18.106	26.194
1765	C	PRO	223	65.741	16.937	25.466
1766	O	PRO	223	65.126	16.370	24.554
1767	CB	PRO	223	65.569	19.408	25.652
1768	CG	PRO	223	66.332	20.150	26.740
1769	CD	PRO	223	66.191	19.310	27.999
1770	N	ASP	224	66.864	16.441	25.966
1771	CA	ASP	224	67.589	15.407	25.205
1772	C	ASP	224	67.038	13.956	25.171
1773	O	ASP	224	67.414	13.283	24.207
1774	CB	ASP	224	69.066	15.400	25.608
1775	CG	ASP	224	69.291	15.200	27.105
1776	OD1	ASP	224	69.769	16.140	27.725
1777	OD2	ASP	224	69.145	14.072	27.554
1778	N	PRO	225	66.204	13.435	26.076
1779	CA	PRO	225	65.596	12.125	25.798
1780	C	PRO	225	64.411	12.131	24.813
1781	O	PRO	225	63.843	11.057	24.581
1782	CB	PRO	225	65.130	11.620	27.129
1783	CG	PRO	225	65.118	12.772	28.116
1784	CD	PRO	225	65.735	13.942	27.375
1785	N	LEU	226	64.031	13.269	24.246
1786	CA	LEU	226	62.873	13.293	23.340
1787	C	LEU	226	63.263	12.821	21.936
1788	O	LEU	226	63.684	13.616	21.086
1789	CB	LEU	226	62.338	14.722	23.261
1790	CG	LEU	226	60.809	14.799	23.223
1791	CD1	LEU	226	60.356	16.247	23.089
1792	CD2	LEU	226	60.190	13.961	22.108
1793	N	ASN	227	63.072	11.536	21.691

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1794	CA	ASN	227	63.324	10.990	20.358
1795	C	ASN	227	62.056	11.083	19.517
1796	O	ASN	227	61.245	10.148	19.479
1797	CB	ASN	227	63.771	9.537	20.490
1798	CG	ASN	227	64.294	9.003	19.157
1799	OD1	ASN	227	63.656	9.153	18.106
1800	ND2	ASN	227	65.447	8.362	19.216
1801	N	ILE	228	62.045	12.076	18.644
1802	CA	ILE	228	60.866	12.346	17.816
1803	C	ILE	228	60.710	11.370	16.644
1804	O	ILE	228	59.569	11.054	16.288
1805	CB	ILE	228	60.985	13.776	17.296
1806	CG1	ILE	228	61.176	14.748	18.455
1807	CG2	ILE	228	59.755	14.169	16.485
1808	CD1	ILE	228	61.294	16.187	17.963
1809	N	CYS	229	61.773	10.665	16.292
1810	CA	CYS	229	61.698	9.714	15.179
1811	C	CYS	229	60.945	8.460	15.611
1812	O	CYS	229	59.923	8.111	15.004
1813	CB	CYS	229	63.120	9.350	14.768
1814	SG	CYS	229	63.276	8.095	13.477
1815	N	VAL	230	61.248	8.016	16.820
1816	CA	VAL	230	60.585	6.839	17.379
1817	C	VAL	230	59.216	7.200	17.952
1818	O	VAL	230	58.291	6.382	17.879
1819	CB	VAL	230	61.486	6.257	18.464
1820	CG1	VAL	230	60.824	5.086	19.182
1821	CG2	VAL	230	62.824	5.829	17.873
1822	N	PHE	231	59.016	8.481	18.217
1823	CA	PHE	231	57.714	8.957	18.686
1824	C	PHE	231	56.700	8.997	17.542
1825	O	PHE	231	55.565	8.542	17.724
1826	CB	PHE	231	57.907	10.361	19.244
1827	CG	PHE	231	57.316	10.574	20.632
1828	CD1	PHE	231	56.927	11.845	21.034
1829	CD2	PHE	231	57.183	9.498	21.500
1830	CE1	PHE	231	56.396	12.039	22.302
1831	CE2	PHE	231	56.651	9.692	22.768
1832	CZ	PHE	231	56.257	10.962	23.168

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1833	N	ILE	232	57.181	9.259	16.335
1834	CA	ILE	232	56.306	9.246	15.160
1835	C	ILE	232	56.021	7.822	14.698
1836	O	ILE	232	54.864	7.515	14.382
1837	CB	ILE	232	56.976	10.037	14.041
1838	CG1	ILE	232	57.103	11.505	14.425
1839	CG2	ILE	232	56.206	9.900	12.731
1840	CD1	ILE	232	57.807	12.305	13.335
1841	N	LEU	233	56.949	6.917	14.972
1842	CA	LEU	233	56.737	5.500	14.641
1843	C	LEU	233	55.797	4.823	15.638
1844	O	LEU	233	54.998	3.964	15.240
1845	CB	LEU	233	58.085	4.790	14.654
1846	CG	LEU	233	59.026	5.370	13.607
1847	CD1	LEU	233	60.428	4.792	13.754
1848	CD2	LEU	233	58.486	5.151	12.198
1849	N	LEU	234	55.692	5.415	16.818
1850	CA	LEU	234	54.766	4.944	17.847
1851	C	LEU	234	53.328	5.375	17.548
1852	O	LEU	234	52.391	4.620	17.839
1853	CB	LEU	234	55.218	5.572	19.163
1854	CG	LEU	234	54.410	5.093	20.362
1855	CD1	LEU	234	54.661	3.612	20.624
1856	CD2	LEU	234	54.759	5.911	21.601
1857	N	LEU	235	53.177	6.411	16.735
1858	CA	LEU	235	51.838	6.899	16.390
1859	C	LEU	235	51.225	6.089	15.249
1860	O	LEU	235	50.003	5.865	15.257
1861	CB	LEU	235	51.938	8.364	15.972
1862	CG	LEU	235	52.564	9.242	17.052
1863	CD1	LEU	235	52.715	10.679	16.563
1864	CD2	LEU	235	51.775	9.200	18.358
1865	N	VAL	236	52.082	5.379	14.526
1866	CA	VAL	236	51.647	4.568	13.379
1867	C	VAL	236	51.007	3.240	13.801
1868	O	VAL	236	50.313	2.609	12.995
1869	CB	VAL	236	52.860	4.276	12.495
1870	CG1	VAL	236	52.441	3.620	11.185
1871	CG2	VAL	236	53.654	5.539	12.194

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1872	N	PHE	237	51.069	2.923	15.087
1873	CA	PHE	237	50.476	1.676	15.572
1874	C	PHE	237	48.951	1.729	15.636
1875	O	PHE	237	48.311	0.671	15.625
1876	CB	PHE	237	51.029	1.379	16.960
1877	CG	PHE	237	52.489	0.943	16.961
1878	CD1	PHE	237	52.944	0.055	15.996
1879	CD2	PHE	237	53.358	1.416	17.934
1880	CE1	PHE	237	54.271	-0.350	15.997
1881	CE2	PHE	237	54.687	1.012	17.934
1882	CZ	PHE	237	55.144	0.130	16.964
1883	N	ILE	238	48.365	2.918	15.672
1884	CA	ILE	238	46.907	2.990	15.561
1885	C	ILE	238	46.511	3.889	14.391
1886	O	ILE	238	45.508	3.634	13.713
1887	CB	ILE	238	46.286	3.483	16.866
1888	CG1	ILE	238	46.826	2.709	18.064
1889	CG2	ILE	238	44.769	3.325	16.809
1890	CD1	ILE	238	46.142	3.132	19.359
1891	N	VAL	239	47.335	4.885	14.107
1892	CA	VAL	239	47.087	5.730	12.934
1893	C	VAL	239	48.095	5.404	11.837
1894	O	VAL	239	49.250	5.853	11.860
1895	CB	VAL	239	47.168	7.196	13.341
1896	CG1	VAL	239	47.005	8.113	12.138
1897	CG2	VAL	239	46.105	7.518	14.379
1898	N	VAL	240	47.640	4.609	10.885
1899	CA	VAL	240	48.519	4.143	9.807
1900	C	VAL	240	48.825	5.260	8.804
1901	O	VAL	240	47.921	5.886	8.238
1902	CB	VAL	240	47.858	2.936	9.139
1903	CG1	VAL	240	46.410	3.217	8.756
1904	CG2	VAL	240	48.653	2.422	7.941
1905	N	LYS	241	50.118	5.500	8.636
1906	CA	LYS	241	50.658	6.524	7.731
1907	C	LYS	241	50.201	7.902	8.173
1908	O	LYS	241	49.214	8.447	7.657
1909	CB	LYS	241	50.226	6.242	6.294
1910	CG	LYS	241	51.261	6.705	5.269

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1911	CD	LYS	241	52.353	5.659	5.020
1912	CE	LYS	241	53.268	5.422	6.218
1913	NZ	LYS	241	54.161	4.284	5.982
1914	N	CYS	242	50.988	8.479	9.063
1915	CA	CYS	242	50.570	9.681	9.776
1916	C	CYS	242	51.610	10.806	9.802
1917	O	CYS	242	52.077	11.190	10.878
1918	CB	CYS	242	50.241	9.230	11.194
1919	SG	CYS	242	51.557	8.399	12.112
1920	N	PHE	243	51.898	11.388	8.650
1921	CA	PHE	243	52.825	12.529	8.619
1922	C	PHE	243	52.698	13.301	7.306
1923	O	PHE	243	52.311	12.740	6.272
1924	CB	PHE	243	54.254	12.009	8.780
1925	CG	PHE	243	55.246	13.038	9.322
1926	CD1	PHE	243	54.862	13.883	10.355
1927	CD2	PHE	243	56.530	13.123	8.799
1928	CE1	PHE	243	55.755	14.822	10.853
1929	CE2	PHE	243	57.423	14.062	9.297
1930	CZ	PHE	243	57.035	14.914	10.323
1931	N	THR	244	52.961	14.595	7.382
1932	CA	THR	244	53.007	15.453	6.198
1933	C	THR	244	54.251	15.132	5.374
1934	O	THR	244	55.141	14.417	5.847
1935	CB	THR	244	53.012	16.906	6.665
1936	OG1	THR	244	52.902	17.764	5.536
1937	CG2	THR	244	54.279	17.268	7.432
1938	N	SER	245	54.198	15.489	4.102
1939	CA	SER	245	55.327	15.247	3.193
1940	C	SER	245	56.631	15.779	3.788
1941	O	SER	245	56.651	16.868	4.375
1942	CB	SER	245	55.049	15.951	1.868
1943	OG	SER	245	54.945	17.344	2.135
1944	N	GLU	246	57.705	15.031	3.566
1945	CA	GLU	246	59.056	15.302	4.100
1946	C	GLU	246	59.052	15.748	5.569
1947	O	GLU	246	59.074	14.881	6.427
1948	CB	GLU	246	59.876	16.257	3.203
1949	CG	GLU	246	59.357	17.684	2.969

Atom No.	Atom Type	Residue	Residue Position	X Coord.	Y Coord.	Z Coord.
1950	CD	GLU	246	58.319	17.802	1.849
1951	OE1	GLU	246	58.152	16.838	1.113
1952	OE2	GLU	246	57.577	18.772	1.885
1953	OXT	GLU	246	59.142	16.946	5.799